

A MICRO-ANALYTIC INVESTIGATION INTO EFL TEACHERS' LANGUAGE TEST  
ITEM REVIEWING INTERACTIONS

A THESIS SUBMITTED TO  
THE GRADUATE SCHOOL OF SOCIAL SCIENCES  
OF  
MIDDLE EAST TECHNICAL UNIVERSITY

BY

HÜMEYRA CAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR  
THE DEGREE OF DOCTOR OF PHILOSOPHY  
IN  
THE DEPARTMENT OF ENGLISH LANGUAGE TEACHING

SEPTEMBER 2020



Approval of the thesis:

**A MICRO-ANALYTIC INVESTIGATION INTO EFL TEACHERS' LANGUAGE  
TEST ITEM REVIEWING INTERACTIONS**

submitted by **HÜMEYRA CAN** in partial fulfillment of the requirements for the degree of  
**Doctor of Philosophy in English Language Teaching, the Graduate School of Social  
Sciences of Middle East Technical University** by,

Prof. Dr. Yaşar KONDAKÇI  
Dean  
Graduate School of Social Sciences

\_\_\_\_\_

Prof. Dr. Çiğdem SAĞIN ŞİMŞEK  
Head of Department  
Foreign Language Education

\_\_\_\_\_

Prof. Dr. Çiler HATİPOĞLU  
Supervisor  
Foreign Language Education

\_\_\_\_\_

**Examining Committee Members:**

Assoc. Prof. Dr. Perihan SAVAŞ (Head of the Examining Committee)  
Middle East Technical University  
Foreign Language Education

\_\_\_\_\_

Prof. Dr. Çiler HATİPOĞLU (Supervisor)  
Middle East Technical University  
Foreign Language Education

\_\_\_\_\_

Assoc. Prof. Dr. A. Cendel KARAMAN  
Middle East Technical University  
Foreign Language Education

\_\_\_\_\_

Assist. Prof. Dr. Ufuk BALAMAN  
Hacettepe University  
Foreign Language Education

\_\_\_\_\_

Assist. Prof. Dr. Hatice ERGÜL  
Hacettepe University  
Foreign Language Education

\_\_\_\_\_





**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

**Name, Last name:** Hümeýra Can

**Signature:**

## ABSTRACT

### A MICRO-ANALYTIC INVESTIGATION INTO EFL TEACHERS' LANGUAGE TEST ITEM REVIEWING INTERACTIONS

CAN, Hümeyra

Ph.D., The Department of English Language Teaching

Supervisor: Prof. Dr. Çiler HATİPOĞLU

September 2020, 578 pages

This study brings an interactional perspective to the construction of syllabus-based language tests and the stage of item reviewing (IR) in particular by using Conversation Analysis (CA). Drawing on a corpus of video-recordings of IR sessions (25 hours) in an English preparatory school at a state university in Turkey, it investigates how EFL teachers review language test items prepared for their students in and through interaction with the item writer who is one of the teachers assigned in the testing office of the school. Following the principles of CA, the initial analysis showed that the interactions mainly focused on the problematic items and they were made up of three recurring phases; (1) problematizing the item (PI), (2) suggesting a change (SC), and (3) reviewing the change (RC). The micro-analysis generated models for the overall structural organization of these emerging phases and revealed recurring interactional actions that led to mutual understanding and joint decision-making about the test item. The teachers were also found to display an orientation to various testing principles in and through talk as they collaboratively reviewed and revised the test items. Additionally, the interactions showed an orientation to the institutional role distribution of the participants although there was clear evidence for their shared responsibility and equal engagement in the IR task. This study is expected to contribute to conversation analytic studies of language testing and teacher-to-teacher institutional talk, and has implications for language test construction, EFL teachers' professional competence, and L2 teacher education.

**Keywords:** item reviewing, language test construction, teacher interaction, conversation analysis, syllabus-based language tests



## ÖZ

### İNGİLİZCEYİ YABANCI DİL OLARAK ÖĞRETEN ÖĞRETMENLERİN DİL SINAV SORULARINI İNCELEME ETKİLEŞİMLERİNİN MİKRO-ÇÖZÜMLEME YOLUYLA İNCELENMESİ

CAN, Hümeýra

Doktora, İngiliz Dili Öğretimi Bölümü

Tez Yöneticisi: Prof. Dr. Çiler HATİPOĞLU

Eylül 2020, 578 sayfa

Bu çalışma Konuşma Çözümlemesi (KÇ) kullanarak müfredata dayalı dil sınavlarının oluşturulmasına, daha belirgin olarak madde inceleme (Mİ) aşamasına etkileşimsel bir bakış açısı getirmektedir. Türkiye’de bir devlet üniversitesinin İngilizce hazırlık okulunda gerçekleşen Mİ oturumlarının video kayıtlarından (25 saat) oluşan bir veritabanı kullanılarak, İngilizceyi yabancı dil olarak öğreten öğretmenlerin öğrencileri için hazırlanmış olan dil sınav sorularını sınavcı ile birlikte etkileşim yoluyla nasıl inceledikleri araştırılmıştır. Sınavcı ise okulun sınav biriminde görevlendirilmiş öğretmenlerden birisidir ve sınavı hazırlayan kişidir. KÇ ilkeleri doğrultusunda verinin ilk analizi etkileşimlerin daha çok problemli maddelere odaklandığını ve bunların tekrarlayan üç aşamadan oluştuğunu göstermiştir; (1) problemi belirleme (PB), (2) değişiklik önerme (DÖ), ve (3) değişikliği gözden geçirme (DG). Yapılan mikro-analizler ortaya çıkan bu üç aşamanın genel yapısal düzeni için modeller sağlamış, sınav maddesine yönelik karşılıklı anlaşılabilirliği ve ortak karar vermeyi sağlayan yaygın etkileşimsel eylemleri ortaya çıkarmıştır. Öğretmenlerin ayrıca sınav maddelerini işbirliği içinde inceleyip düzeltirken konuşma yoluyla çeşitli sınav hazırlama ilkelerine yönelim gösterdikleri bulunmuştur. Buna ek olarak, katılımcıların Mİ görevinde ortak sorumluluk ve eşit katılımını gösteren açık kanıtlar olmasına rağmen etkileşimler katılımcıların kurumsal rol dağılımına bir yönelim sergilemiştir. Bu çalışmanın dil sınavlarına ve öğretmenler arasında gerçekleşen kurumsal konuşmalara yönelik yapılmış



KÇ alıřmalarına katkı saęlaması beklenmektedir. Ayrıca bu alıřma dil sınavı oluřturma alanı, İngilizceyi yabancı dil olarak öreten öretmenlerin mesleki yeterlilikleri ve ikinci dil öretmen yetiřtirme programları için uygulamalar sunmaktadır.

**Anahtar Kelimeler:** madde inceleme, dil sınavı oluřturma, öretmen etkileřimi, konuřma özümlemesi, müfredata dayalı dil sınavları



To My Beloved Parents

*Emine and Yunus Can*

## ACKNOWLEDGMENTS

The long and challenging journey that I took in writing up this dissertation could not have been successfully completed without the heart-warming company of the following people. I would like to express my gratitude to all those who have made this dissertation possible.

First of all, I would like to wholeheartedly thank my supervisor Prof. Dr. Çiler Hatipoğlu for being a great source of motivation and inspiration throughout the process of writing up this dissertation with her invaluable feedback, guidance, intellectual as well as moral support and encouragement. It has been an immense privilege and honour to be her student and to have had the chance to work with her from the very beginning of my undergraduate education and all throughout my graduate studies. Her expertise and experience in the field as well as her great enthusiasm for research has made her a role model in my academic career, contributing to my development as a researcher. I would also like to express my deep sense of gratitude to the jury members of my thesis advisory committee, Assoc. Prof. Dr. A. Cendel Karaman and Assist. Prof. Dr. Ufuk Balaman, and to the previous jury members, Assoc. Prof. Dr. Olcay Sert and Prof. Dr. İsmail Hakkı Erten whom I remember with respect, for their constructive feedback, expert opinion, and helpful suggestions, which enabled me to approach my research critically and to keep on track. I would like to extend my gratitude to the other jury members, Assoc. Prof. Dr. Perihan Savaş and Assist. Prof. Dr. Hatice Ergül for their valuable time and significant contributions to this dissertation.

My special thanks go to the leading members of HUMAN Research Centre, Assoc. Prof. Dr. Olcay Sert, Assist. Prof. Dr. Ufuk Balaman, and Assist. Prof. Dr. Nilüfer Can Daşkın for lending me a helping hand whenever I wanted their guidance on my data analysis, giving me the opportunity to present my data in data sessions and for helping me develop a better understanding of Conversation Analysis. I would also like to take this opportunity to express my appreciation to the other researchers who participated in the data sessions for sharing their points of analysis on the segments of my data. This dissertation would not have been possible without the instructors/testers and colleagues who helped me tremendously in the data collection process. I owe a depth of gratitude to them for their kindness and collaboration during the challenging task of recording my data.

My heartfelt thanks go to my friends and colleagues Çiğdem Kayıhan Aslan, Ece Er, Mustafa Öztürk, Şenay Saraç, Yusuf Emre Özdemir, and Zehra Daşkın for their emotional support, encouragement, and belief in me and most importantly, for always being there ready to meet and talk just when I needed. I am deeply grateful to them for their sincere friendship and company. I would also like to express my appreciation to my friend and PhD fellow Fatma Gümüşok for her great company as we had study hours together and motivated each other whenever we felt less productive and needed a change.

No words can ever be strong enough to express my indebtedness to my family. My wholehearted and special thanks go to my beloved parents Emine and Yunus Can, to whom this dissertation is dedicated, for their limitless support, patience, well-wishes, prayers, unconditional love and affection, and genuine belief in me. We have walked together on this journey, sharing every moment and the bitter-sweet memories on the way. I am eternally grateful to them for their tolerance for long working hours, great sacrifices just so I could work effectively and successfully complete my study, and for the opportunities they have created for me to pursue my goals. My heartfelt thanks go to my dear sisters Nilüfer and V. Betül for doing everything they can to keep me going and for always being by my side as my besties, friends, companions, colleagues, and advisors. They were always the first ones to talk to whenever I felt stressed, hopeless, and stuck in the process. Their continual intellectual and emotional support made this challenging journey much more meaningful, manageable, and less painful. As we started our academic career together, they were the ones who could best understand what I was going through. My youngest sister V. Betül's time management and writing tips and her effective solutions to my problems made my life a lot easier. My twin sister Nilüfer's reviews of my drafts and her valuable comments, suggestions, and insights as an expert in the same field contributed significantly to the development of this dissertation and made me feel much more confident in what I was doing. I am deeply indebted to her for always being available to brainstorm ideas, answer my questions patiently and clear up my confusion. I would also like to express my appreciation to my elder sister Gülistan and my brother Fatih for their support and care during the ups and downs of this journey. I should further send my love to my little niece Beyza for filling me with joy in the stressful times.

Finally, I would like to offer my sincere thanks to everyone who in any respect supported me during the completion of this dissertation.

## TABLE OF CONTENTS

|                                                                                                         |      |
|---------------------------------------------------------------------------------------------------------|------|
| PLAGIARISM .....                                                                                        | iii  |
| ABSTRACT.....                                                                                           | iv   |
| ÖZ.....                                                                                                 | vi   |
| DEDICATION.....                                                                                         | viii |
| ACKNOWLEDGMENTS.....                                                                                    | ix   |
| TABLE OF CONTENTS.....                                                                                  | xi   |
| LIST OF TABLES.....                                                                                     | xiv  |
| LIST OF FIGURES.....                                                                                    | xvi  |
| LIST OF ABBREVIATIONS.....                                                                              | xvii |
| CHAPTERS                                                                                                |      |
| 1. INTRODUCTION.....                                                                                    | 1    |
| 1.0. Introduction.....                                                                                  | 1    |
| 1.1. Background to the Study.....                                                                       | 1    |
| 1.2. Purpose and Significance of the Study.....                                                         | 6    |
| 1.3. Research Questions.....                                                                            | 14   |
| 1.4. Terminology Used.....                                                                              | 14   |
| 2. REVIEW OF LITERATURE.....                                                                            | 17   |
| 2.0. Introduction.....                                                                                  | 17   |
| 2.1. Language Test Construction.....                                                                    | 17   |
| 2.2. Syllabus-based Language Tests.....                                                                 | 21   |
| 2.3. Item Reviewing (IR) Practice in Language Test Construction.....                                    | 22   |
| 2.4. Interaction and Teacher Involvement in Item Reviewing Practice.....                                | 27   |
| 2.5. Studies on Interaction in Language Test Construction.....                                          | 32   |
| 2.6. Conversation Analysis (CA).....                                                                    | 39   |
| 2.6.1. Institutional Talk.....                                                                          | 60   |
| 2.6.2. Conversation Analytic Studies on Teacher-to-teacher Collaborative<br>Workplace Interactions..... | 71   |
| 2.6.3. Conversation Analytic Studies in Language Testing/Assessment.....                                | 77   |
| 3. METHOD OF RESEARCH.....                                                                              | 80   |
| 3.0. Introduction.....                                                                                  | 80   |

|                                                                                                                  |     |
|------------------------------------------------------------------------------------------------------------------|-----|
| 3.1. Conversation Analysis as a Research Methodology.....                                                        | 80  |
| 3.2. Research Context.....                                                                                       | 84  |
| 3.3. Participants.....                                                                                           | 89  |
| 3.4. Data Collection Procedures.....                                                                             | 96  |
| 3.5. Data Transcription and Analysis Procedures.....                                                             | 104 |
| 4. DATA ANALYSIS AND FINDINGS.....                                                                               | 118 |
| 4.0. Introduction.....                                                                                           | 118 |
| 4.1. Phase 1. Problematizing the Item (PI).....                                                                  | 118 |
| 4.1.1. Summary of Interactional Patterns in Phase PI.....                                                        | 219 |
| 4.2. Phase 2. Suggesting a Change (SC).....                                                                      | 243 |
| 4.2.1. Summary of Interactional Patterns in Phase SC.....                                                        | 346 |
| 4.3. Phase 3. Reviewing the Change (RC).....                                                                     | 378 |
| 4.3.1. Summary of Interactional Patterns in Phase RC.....                                                        | 432 |
| 5. DISCUSSION.....                                                                                               | 450 |
| 5.0. Introduction.....                                                                                           | 450 |
| 5.1. Overall Structural Organization of Language Test IR Interactions in the<br>Case of Problematized Items..... | 450 |
| 5.1.1. The Phases of Language Test IR Interactions.....                                                          | 450 |
| 5.1.2. The Sequential Environment of Each Phase.....                                                             | 455 |
| 5.1.2.1. Problematizing the item (PI).....                                                                       | 455 |
| 5.1.2.2. Suggesting a Change (SC).....                                                                           | 470 |
| 5.1.2.3. Reviewing the Change (RC).....                                                                          | 483 |
| 5.2. Testing Principles Oriented to Throughout the Phases of Language Test IR<br>Interactions.....               | 488 |
| 5.3. Institutional Role Distribution of Participants in the Phases of Language Test<br>IR Interactions.....      | 496 |
| 5.4. Implications for Language Test Construction.....                                                            | 503 |
| 5.5. Implications for EFL Teachers' Professional Competence.....                                                 | 508 |
| 5.6. Implications for L2 Teacher Education.....                                                                  | 514 |
| 6. CONCLUSION.....                                                                                               | 518 |
| 6.0. Introduction.....                                                                                           | 518 |
| 6.1. Conclusions and Implications.....                                                                           | 518 |
| 6.2. Limitations.....                                                                                            | 527 |
| 6.3. Further Research.....                                                                                       | 528 |

|                                                          |     |
|----------------------------------------------------------|-----|
| REFERENCES.....                                          | 530 |
| APPENDICES                                               |     |
| A. BACKGROUND QUESTIONNAIRE.....                         | 551 |
| B. JEFFERSON TRANSCRIPTION CONVENTIONS.....              | 554 |
| C. MEANINGS OF FLOWCHART SYMBOLS USED IN FIGURES.....    | 556 |
| D. APPROVAL OF METU HUMAN SUBJECTS ETHICS COMMITTEE..... | 557 |
| E. INFORMED CONSENT FORM.....                            | 558 |
| F. CURRICULUM VITAE.....                                 | 559 |
| G. TURKISH SUMMARY / TRKE ZET.....                    | 561 |
| H. THESIS PERMISSION FORM / TEZ İZİN FORMU.....          | 578 |



## LIST OF TABLES

|                                                                                                                        |     |
|------------------------------------------------------------------------------------------------------------------------|-----|
| Table 1 Teachers' Demographic Information.....                                                                         | 90  |
| Table 2 Testers' Demographic Information.....                                                                          | 93  |
| Table 3 Details about the Language Test IR Data.....                                                                   | 99  |
| Table 4 Test Task Types Used in the Reviewed Tests in the Data.....                                                    | 102 |
| Table 5 Test Task Types According to the Sections in the Reviewed Tests in the Data..                                  | 103 |
| Table 6 Interactional Resources for Initiating Preliminary Inquiry into the Problem.....                               | 221 |
| Table 7 Interactional Resources for Projecting the Problem in the Item.....                                            | 223 |
| Table 8 Summary of Accounts and Evidence Presented for the Problems in the Items...                                    | 225 |
| Table 9 Interactional Resources for Mitigating the Problem in the Item.....                                            | 228 |
| Table 10 Summary of Challenges Presented to the Announced Problems in the Items<br>and Their Outcomes.....             | 231 |
| Table 11 Interactional Resources for Displaying Mutual Agreement through a<br>Collaborative Process in Phase PI.....   | 234 |
| Table 12 Interactional Indicators for Tester Roles and Responsibilities in Phase PI.....                               | 236 |
| Table 13 Interactional Indicators for Reviewer Roles and Responsibilities in Phase PI..                                | 238 |
| Table 14 Summary of Suggestions and Decided Changes for the Problems in the<br>Items.....                              | 348 |
| Table 15 Interactional Resources for Downgrading or Displaying Certainty towards a<br>Suggestion.....                  | 354 |
| Table 16 Summary of Accounts and Evidence Presented for the Suggestions.....                                           | 359 |
| Table 17 Summary of Challenges Presented to the Suggestions and Their Outcomes...                                      | 367 |
| Table 18 Interactional Resources for Displaying Mutual Agreement through a<br>Collaborative Process in Phase SC.....   | 371 |
| Table 19 Interactional Indicators for Tester and Reviewer Roles and Responsibilities<br>in Phase SC.....               | 373 |
| Table 20 Summary of Challenges Presented to the Change and Their Outcomes.....                                         | 435 |
| Table 21 Summary of Suggestions and Decided Changes for the Problems in the<br><i>Revised</i> Language Test Items..... | 439 |
| Table 22 Interactional Resources for Displaying Mutual Agreement and<br>Understanding in Phase RC.....                 | 444 |



|                                                                                                                  |     |
|------------------------------------------------------------------------------------------------------------------|-----|
| Table 23 Summary of Interactional Indicators for Tester and Reviewer Roles and Responsibilities in Phase RC..... | 445 |
| Table 24 Recurring Interactional Actions in Phase PI.....                                                        | 456 |
| Table 25 Recurring Interactional Actions in Phase SC.....                                                        | 471 |
| Table 26 Recurring Interactional Actions in Phase RC.....                                                        | 484 |
| Table 27 Testing Principles Oriented to in Language Test IR Interactions.....                                    | 489 |
| Table 28 Tester and Reviewer Roles and Responsibilities Made Relevant in Language Test IR Interactions.....      | 497 |



## LIST OF FIGURES

|                                                                                                                      |     |
|----------------------------------------------------------------------------------------------------------------------|-----|
| Figure 1 Results of the Unmotivated Examination and Initial Analysis of the Data.....                                | 113 |
| Figure 2 Procedure Followed in the Data Analysis.....                                                                | 117 |
| Figure 3 Overall Structural Organization of Phase PI.....                                                            | 219 |
| Figure 4 Overall Structural Organization of Phase SC.....                                                            | 346 |
| Figure 5 Overall Structural Organization of Phase RC.....                                                            | 432 |
| Figure 6 Overall Structural Organization of Language Test IR interactions in the Case<br>of Problematized Items..... | 451 |
| Figure 7 The Model Emerging from Language Test IR Interactions.....                                                  | 515 |

## LIST OF ABBREVIATIONS

|      |                                               |
|------|-----------------------------------------------|
| CA   | Conversation Analysis                         |
| DIU  | Designedly Incomplete Utterance               |
| EFL  | English as a Foreign Language                 |
| FPP  | First Pair Part                               |
| IR   | Item Reviewing                                |
| LAL  | Language Assessment Literacy                  |
| L2   | English as foreign/second/additional language |
| NMP  | Non-minimal Post-expansion                    |
| PI   | Problematizing the Item                       |
| RC   | Reviewing the Change                          |
| SC   | Suggesting a Change                           |
| SLTE | Second Language Teacher Education             |
| SPP  | Second Pair Part                              |
| TCU  | Turn Constructional Unit                      |
| TLU  | Target Language Use                           |
| TRP  | Transition Relevance Place                    |

## **CHAPTER 1**

### **INTRODUCTION**

“When we grow up, we ‘become human’ in and through talk and much of our social life is in fact enacted as and in talk. Think of education, medical care, politics, commerce, and, indeed, science; none of these crucial social activities could do without talk in some way. But even in situations which are not generally seen as ‘important’, like chatting during a break, we manifest ourselves and perceive others largely through our talking together.”

— (Ten Have, 2007, p.10)

#### **1.0. Introduction**

In this chapter, background to the study, the purpose and significance of the study, the research questions to be investigated and the terminology used in the study are presented and discussed.

#### **1.1. Background to the Study**

Language test construction has always been an important professional practice in language teaching programs for several reasons. Firstly, language tests are widely used to serve various purposes such as placement, progress, diagnostic, achievement and proficiency (Alderson et al., 1995). They aim to (1) measure overall language proficiency, (2) find out the extent to which students have achieved the objectives of a course of study, (3) diagnose students’ strengths and weaknesses, (4) identify what they know and do not know and (5) discover their level of proficiency to be able to place them into the right classes in a language program (Bachman, 1990; Douglas, 2010; Henning, 1987; Hughes, 1989). According to Bachman (1990), tests are valuable in that they lead to the elicitation of the specific kinds of behaviour that the test user can interpret as evidence of the abilities or attributes in question. Secondly, language tests provide valuable sources of information about the effectiveness of learning, instruction and the teaching program, which enables teachers to make the necessary revisions in the structure and content of future lessons or in

their approaches to teaching, and which also allows students to see how they are doing and what they can do to study more effectively (Bachman, 1990; Brookhart, 1999; Douglas, 2010; Heaton, 1990; Henning, 1987). Teachers also use tests to motivate their students to study by giving them the opportunity to demonstrate their ability to perform certain tasks in the target language (Fulcher, 2010; Heaton, 1990). Thirdly, language tests have a gate keeping role and based on the test results, life-changing or high stakes decisions can be made about individuals within their educational context (Bachman, 1990; Douglas, 2010; McNamara, 2004; McNamara & Roever, 2006). For example, tests can determine whether the students have reached the proficiency level required to start their departments at university. Finally, as Douglas (2010) points out, language tests ensure fairness and standardisation by giving students an equal opportunity to show their performance in the language that they have learned and allowing teachers to judge their students' progress against a certain performance criteria. They also provide teachers with more confidence by confirming their own assessments and decisions about students' progress (Douglas, 2010).

Although there is no one 'best' test for any given situation, to provide all these advantages and functions, it is necessary that the tests are of high quality, serving the intended purposes and that their results are interpreted appropriately. The fact that the language tests can be used for various purposes, playing vital roles for educational settings makes language test construction a critical process. Even though it is not possible to control and eliminate all the irrelevant factors that influence language test performance, their effects can be minimized while maximizing the effects of the target ability by following a particular set of test development procedures (Bachman, 1990). This would increase the validity of the interpretation and use of a test score as an indicator of a certain ability.

One such test development procedure consists of the practice of language test item reviewing (IR). The inspiration and motivation to carry out this study first came from the researcher's experience as a tester for five years as well as a regular language test item reviewer at an English preparatory school in a higher education setting in Turkey, which was the research context for this study. In the majority of English preparatory schools at universities in Turkey, there are testing units where language teachers already working as instructors in the prep school are assigned to work as testing unit supervisors and where they go through similar processes in preparing their language tests. Those testers do not have many hours of teaching and thus, they get the teachers who have classes and who actually

teach the content of the prepared tests to review the tests collaboratively to make sure that the exam tests what and how it is supposed to test. This practice is generally referred to as proofreading. It is one of the most common and important practices at those testing units and is greatly valued by the testers, who might or might not have testing experience in that institution. In this respect, IR sessions are generally viewed as an important part of the test validation process and as learning opportunities for the testers as well as the teachers although their effectiveness can change depending on the reviewers and the quality of their interaction with the testers. Therefore, it is possible to hear testers making comments such as the following:

*I get happy when more and more mistakes and problems are found in the tests during the test reviewing sessions because this prevents serious problems that we might encounter after the exam is administered.*

*When I first began working as a testing unit supervisor, I only had the theoretical knowledge I gained in my undergraduate studies and I was really nervous; however, when I got involved in the testing process, I realized that it was a dynamic activity and it was mainly the test reviewing sessions where I slowly developed a testing perspective as I interacted with the teachers.*

The teachers are also encouraged to support the testers by accepting to review their tests as the product they create is the product of the institution and they try to achieve a common goal. Furthermore, since IR is the step that gives the final shape to the exams, language teachers also share the responsibility and it requires a careful analysis of the exam papers in line with the materials and instructional events and thus, critical thinking as well as problem solving skills.

All these show that when language teachers start their profession, they are not only expected to teach but also get involved in the test construction process either by being selected to work at a testing unit or reviewing exam papers prepared by the testers. This means that they have multiple roles and need to be prepared for such a practice before they graduate. However, in order to gain IR skill, the processes of language test IR need to be discovered and shared with student teachers. As Madraso (1993) makes it clear, “proofreading is a necessary skill that is much talked about but rarely taught. In fact, most of us teachers have not been taught specific proofreading skills” (p.32). This study, therefore, brings into view the practice of IR as an important part of language test co-construction by systematically describing it as an interactional activity via Conversation Analysis (CA). It focuses on the interactional processes involved in IR and on what happens when teachers together review

and revise language tests in and through talk. In this way, the scrutinized overall structural organization of the IR sessions and the interactional actions that lead to joint decision-making as well as changes and improvements in the test items can be used as part of a framework for the development of practitioners' IR skill and interactional competence for language test construction.

Language test construction has generally been described in terms of a number of set of stages and models, most of which also include item reviewing and editing as one of the critical stages (e.g., Bachman & Palmer, 1996; Carr, 2011; Downing, 2006; Fulcher, 2010; Fulcher & Davidson, 2007; Hughes, 1989; Izard, 2005; Milanovic, 2002; Mislevy et al., 2002) and interaction with stakeholders such as teachers has been highlighted especially at this stage (e.g., Carr, 2011; Fulcher, 2010; Spaan, 2007). It has been well recognized that classroom test construction is not merely an individual activity whereby the testers or teachers get the materials covered in class and prepare the tests in line with the content and tasks in those materials as well as the testing principles. As a socially constructed practice that ultimately affects a large group of people, it goes far beyond this and requires not just the testing skills, knowledge and principles as well as technical or statistical procedures but also interaction with stakeholders (e.g., teachers) within a specific institutional context as a way to better address the dynamic interplay between teaching and testing (Bachman & Palmer, 1996; Fulcher, 2010; Fulcher & Davidson, 2007; Hughes, 1989; Izard, 2005; Mathew, 2004; Shohamy, 2007). In this respect, it is acknowledged that the real world of testing can be much messier than expected because test developers often work in a complex context with various constraints and have to interact with stakeholders, who might have different interests and perspectives (Buck, 2009). To produce good testing instruments, test constructors, thus, need to be able to apply their professional skills in creative ways to novel situations, which requires negotiation, compromise and concession (Buck, 2009).

In the literature; however, the descriptions of language test construction have mostly been general and vague lacking concrete and empirical details with too much focus on the theoretical aspects and principles. The determined procedures and texts on language testing have mainly addressed large-scale testing, resulting in the likelihood of mixing technical large-scale definitions with everyday usages of common testing terms (Fulcher & Davidson, 2007). Thus, they might not always be relevant for in-house language test construction as the issue of validity is approached from a different perspective with less consideration for the

needs of students and classroom context (Fulcher, 2010; Tan & Turner, 2015). This is also acknowledged by Brindley (2001) who points out that the language testing literature is heavily concerned with theoretical and statistical aspects of test validity and reliability and with techniques drawn from large-scale standardized test construction and validation, and tends to be irrelevant to the daily assessment practices of teachers.

In texts on language testing, IR practice in particular has generally been described in terms of sub-processes (Fulcher & Davidson, 2007; Fulcher, 2010) and checklists have been suggested to guide those processes, but there has been no focus on their interactional nature (Bachman & Palmer, 1996; Brown, 1996; Brown, 2001; Izard, 2005; Weir, 2005). Besides, even though there are a few current studies that deal with teacher feedback and the role of interaction with stakeholders at different stages of language test construction (See Section 2.5), they mainly focus on large-scale tests and use primarily interviews and focus groups. Therefore, the micro-details of the actual interactions between teachers and testers in reviewing language tests have not been examined. Shin (2012) pinpoints that there is a need for “micro-level studies of specific facets of test development” and of those facets, the actual process of item writing specifically remains to be an area for further study (p.245) although it is one of the critical phases of test development (Kim et al., 2010). With regard to this, Davidson and Lynch (2002) and So (2014), who emphasize teamwork in language test development, revealed that in language testing literature, there is very little information and guidance on how people write tests in groups and how they incorporate stakeholder feedback in constructing tests. The reason for this is that many language testing books published in recent decades have explained the item writing process in terms of a tightly organized set of guidelines, procedures and validity arguments and have given more space for the development of test specifications, neglecting the issue of how test items are written, revised and developed in the real world, how applicable the suggested guidelines are (e.g., writing test items from test specifications), and whether there are any sociocultural factors affecting the item writing process (Kim et al., 2010; Shin, 2012). In relation to this, Shin (2012) states that “item writing needs to be understood as a social practice” (p.243) as “different stakeholders, different purposes, different cultures of testing can lead to different thinking” (p.245) and hence, item writing can be realized in ways that are different from the “taken-for-granted” procedures explained in testing literature. This requires item writing to be examined in greater detail through qualitative research methods. Therefore, for a better understanding of testing in local contexts, which future teachers are more likely to encounter



and get engaged in and to be more helpful for practitioners, there is a need to go beyond highly abstract theoretical concepts and a need for a process-oriented study of language test construction as it is realized in its own local context as part of teachers' day-to-day work.

To summarize, although books and texts on language testing present IR practice as a critical stage in their step-by-step descriptions of language test construction while at the same time indicating the iterative, dynamic and complex nature of test construction process and emphasize the role of interaction with stakeholders especially at the stage of item reviewing and editing (Bachman & Palmer, 1996; Buck, 2009; Carr, 2011; Fulcher & Davidson, 2007; Kim et al., 2010), the interactional resources that participants enact during talk-in-interaction while reviewing language tests and the interactional architecture of IR interactions have not been studied. More specifically, how actually in-house language tests are conjointly constructed in and through talk, how testers and teachers achieve and manage intersubjectivity as they collaboratively review tests and how they orient to testing principles in and through interaction is an unexplored area of study. In other words, the micro-details of interactions that take place between teachers in the process of reviewing language tests for a specific content area and a group of students are unknown. Therefore, this study brings an interactional perspective to the co-construction of syllabus-based language tests by using CA and providing a micro-analytic description of the recurring phases and interactional features that make up language test IR sessions among EFL teachers in a particular institutional context.

## **1.2. Purpose and Significance of the Study**

This study provides a systematic description of EFL teachers' reviewing of syllabus-based language test items as an interactional activity by using CA. In doing this, this study aims to;

- (1) investigate the interactional processes of language test IR practice as instantiated by the teachers through talk,
- (2) identify the phases of language test IR interaction and develop a model for its overall structural organization,
- (3) reveal the recurring interactional actions that could be valuable for the description of practitioners' language test IR competence,
- (4) show how participants orient to testing principles and deal with them in and through interaction, and

(5) explain what specific roles and responsibilities the participants display in their turns-at-talk as reviewers and testers.

Through a micro-analytic investigation of syllabus-based language test IR interactions between EFL teachers in an educational context, this study is expected to;

(1) develop a better understanding of the IR stage in language test construction by demonstrating in practice how the participants achieve the goal of IR practice in and through interaction, how they establish joint decision-making and mutual understanding about the test items and what consequences the IR practice has for test items,

(2) highlight the importance of interaction as well as teamwork and teacher involvement in the reviewing and co-construction of language tests,

(3) turn the attention to the co-construction of syllabus-based language tests in local contexts, which have been neglected as compared to large-scale testing,

(4) contribute to qualitative studies in the field of language assessment and more specifically, to conversation analytic language testing studies, which have mainly dealt with language proficiency interactions in speaking tests, and

(5) provide the conversation analytic investigation of a different institutional talk in educational setting (i.e., teacher-to-teacher interactions in testing units in a higher education setting), going beyond classroom interaction and hence, contribute to studies on institutional conversation analysis and workplace interactions.

The current study focuses on a specific EFL testing context in a particular institution at tertiary level but it is hoped that the findings and general conclusions will be relevant and transferable to the practice of syllabus-based language test construction in other language teaching programs in Turkey and around the world where tests are locally prepared and classroom teachers together review and co-construct language tests.

This study has a number of implications. First of all, this study is expected to contribute to language test construction and the stage of IR in particular by providing a data-driven interactional model for the phases of language test IR practice. In this way, the model can suggest the interactional actions and resources that are necessary for reviewing, editing and revising tests collaboratively. More specifically, this study is expected to present a more vivid picture of IR practice for syllabus-based language tests by showing the micro-analytic details of how test items are problematized, suggestions are provided to change and improve the item, changes are reviewed, and how mutual agreement and understanding is reached in

the group during the process of finalizing the item in and through interaction in the IR meetings with teachers. This would help understand the relation between test construction and interaction, which has been emphasized in the testing literature (e.g., Bachman & Palmer, 1996; Fulcher, 2010; Fulcher & Davidson, 2007; Hughes, 1989; Izard, 2005; Mathew, 2004; Shohamy, 2007) but which has not been examined in detail. Furthermore, the micro-analytic investigation of language test IR interactions in this study will show one of the qualitative procedures to validate test items in the process of locally developing language tests and hence, demonstrate how validity evidence can be obtained from the stakeholders through qualitative item reviewing in and through interaction (Downing, 2006; Weir, 2005; Wolf et al., 2008).

Secondly, this study has implications for EFL teachers' professional competence and second language teacher education programs by providing a micro-analytic description of the phases and the interactional actions required to develop the interactional competence that is necessary for reviewing and revising language tests collaboratively. Such a description of language test IR interactions is important considering teachers' roles as testers, their testing needs and the necessity of teaching how to review language test items systematically in and through talk in teacher education programs (i.e., both pre-service and in-service) as it is a practice which future teachers will be highly likely to engage in.

(1) *Teachers' roles as testers:*

Teachers' testing and assessment responsibilities have increased to a great extent due to accountability issues, the washback of tests on teaching and the recognition of assessment for learning as an essential component of classroom practice (Fulcher, 2012). According to Stiggins (1999), overall "the typical teacher can spend as much as a third to a half of his or her professional time involved in assessment-related activities" and "competence is required to do this job well" (p.23). When they start their profession, language teachers are not only expected to teach but also prepare and administer various tests to different groups of students and hence, second language assessment constitutes an important part of language teachers' expertise (O'Loughlin, 2006). This is true especially for exam-oriented education systems. In Turkey, for example, at every level, there are compulsory class evaluations, most of which are prepared by classroom teachers and teachers' role as testers is one of their most important roles (Hatipoğlu, 2015). Therefore, teachers in Turkey "are expected to have a

solid knowledge of ELTE [English Language Testing and Evaluation], which will help them fit quickly into the system and deliver successful programmes” (Hatipoğlu, 2010, p.43).

Besides having an understanding of the conceptual bases of assessment and testing, teachers should also be able to relate such knowledge to their professional practice in their specific context (Scarino, 2013). As Brindley (1997) further indicates, due to the increasing demands and pressures to provide explicit information on program activities and outcomes, in many educational institutions, teachers find themselves “in the position of having to develop tools and procedures for monitoring, recording, and assessing learners’ progress and achievement in the classroom on a more systematic and formal basis” (p.37). To be able to carry out formal assessments, they need a wide range of skills such as designing classroom tests and assessment tasks and evaluating their quality (Brindley, 1997). Even if teachers are not directly responsible for constructing the tests themselves, they need to understand how tests should be constructed and what constitutes good testing practice so that they can gain awareness about the assessment process, evaluate the claims made for tests by different examining authorities, and from among a range of available tests, select the one that best suits their purpose and context or edit and moderate tests prepared within their institution (Alderson et al., 1995). Therefore, teachers are required not only to design and develop quality assessment exercises but also to learn what can go wrong with them and how to minimize errors and avoid problems (Stiggins, 1999). In other words, as expected from assessment literates, they should be able to “know the difference between sound and unsound assessment”, being sensitive to the potential negative consequences of inaccurate assessment and are not “intimidated by the sometimes mysterious and always daunting technical world of assessment” (Stiggins, 1995, p.240). In addition, they should know what to do when they encounter a contradictory view in teacher interactions on assessment issues, view disagreement and conflict not as a threat to validity but as a source of validity and be able to resolve conflict by exchanging advice with others in appropriate and structured ways (Davison, 2004). English language teachers also “need to develop critical strategies to examine the uses and consequences of English language tests, to control the power of tests and minimize their detrimental impact”, which requires teachers to become socially responsible and reflexive about the uses of language tests (Shohamy, 2007, p.530).

Language teachers’ ability to evaluate assessment tools within their social context with regard to their purposes is also pinpointed by Inbar-Lourie (2008) in her definition of

assessment literacy, which states “having the capacity to ask and answer critical questions about the purpose for assessment, about the fitness of the tool being used, about testing conditions, and about what is going to happen on the basis of the results” (p.389). In a similar vein, Fulcher (2012) summarizes the language assessment literacy required of teachers as: (1) the knowledge, skills and abilities that are necessary to work with standardized or classroom based tests, which includes designing, developing, and evaluating them, (2) teachers’ familiarity with and awareness of the test processes, principles and concepts, and (3) teachers’ ability to consider assessment within wider historical, social, political and philosophical frameworks to better understand and evaluate the assessment practices. This means that the IR skill and interactional competence that is proposed in this study for language test construction is an important part of EFL teachers’ assessment literacy as well as their institutional responsibilities. They need to develop their ability to evaluate and revise language tests collaboratively in and through interaction with their colleagues to be able to fulfil their testing responsibilities successfully in their professional contexts.

*(2) Teachers’ testing needs and teacher education programs:*

There are studies showing that in pre-service teacher education programs, there is a need to devote more time to (a) constructing tests collaboratively through interactions between student teachers, which provide possibilities to negotiate meaning, (b) critically evaluating, reviewing and discussing items from various published exams as well as from actual classroom-based tests prepared by student teachers, and (c) questioning their validity while at the same time reading texts on test development; and there is also a need to form a professional community of practice and to present student teachers with the real challenges of day-to-day testing practices in specific language education contexts (e.g., Hatipoğlu, 2010; Kleinsasser, 2005; O’Loughlin, 2006). In a study on the testing needs of teachers, Fulcher (2012) found that teachers need a practical how-to guidance that is not necessarily prescriptive in nature but that explains the ongoing test design process within a social context. According to him, such a how-to guidance should include illustrations and practical examples drawn from a range of sources and countries for both classroom and large-scale testing as well as activities that can be reasonably undertaken considering the constraints, resources and contextual issues. Furthermore, pre-service or novice teachers are required to be acquainted with the skills and competencies that they can use wherever they are to meet the demands and challenges of their specific contexts. As Davidson and Lynch (2002)

suggest, rather than offering preconceived notions about testing, an attempt should be made to help teachers acquire tools that are necessary to write their own recipes that reflect the unique needs of their setting. The IR skill and interactional competence that are discussed in relation to test construction in this study can be considered to be part of those tools, which are necessary for teachers to co-construct quality tests for their purposes in line with the needs of their specific contexts.

Student teachers may find it difficult to develop assessment materials, tasks and items and to put their theoretical knowledge into practice since it is not the typical process that they go through in a real classroom and institutional context (Kleinsasser, 2005) and since the high level of abstraction around key theoretical concepts such as validity and reliability makes second language assessment “a notoriously difficult domain of knowledge for students in second language teacher education programs” (O’Loughlin, 2006, p.71). As for the current language testing textbooks, although they provide practical advice on test construction, they still have a tendency to bring forward statistical models that are complicated for newcomers to understand and are written mainly for language testers and experts (Davidson & Lynch, 2002). Therefore, they seem to be too technical and specialized for language educators (Brindley, 2001; Taylor, 2009).

Since teachers face time and resource constraints in their testing practices, it is important for testing educators to develop flexible approaches to their assessment practices and to discuss the social context of testing (Brindley, 2001). With regard to this, Brindley (2001) further indicates that professional development in assessment should not be just about demonstrating its technical sides because “assessment does not happen within a vacuum” and “teachers need the opportunity to see how assessment within their own organisation fits into the broader social context” (p.131). Of the eight dimensions of language assessment literacy proposed by Taylor (2013), she also includes sociocultural values, local practices and personal beliefs besides dimensions such as knowledge of theory, technical skills, principles and concepts and emphasizes engagement with test stakeholder groups in developing a test that meets their specific needs. As Hatipoğlu (2015) further states, assessment is context dependent, which means that the needs related to assessment change depending on the local context and that there is not a fix set of rules and procedures for implementing assessment. This means that pre-service teachers need to be familiarized with the actual testing processes that take place at educational institutions both in and outside

classrooms and to be shown what they will be doing in relation to testing once they are on the job. This study aims to do this by presenting the micro-analytic details of language test IR interactions between teachers and testers in a local institutional context, which future teachers will be likely to engage in.

*(3) Teaching how to review language test items systematically in and through talk in teacher education programs:*

Kelly et al. (2004) introduce “peer review” as a strategy to be implemented in language teacher education programs to develop the skills of team-working, communicative ability, cooperation and practical analysis as well as the ability to provide constructive feedback with a critical approach and to value a sense of professional community spirit, which can all be important for working professionally. This suggests giving teacher trainees guidelines about how to structure peer review and how to provide critical feedback (Kelly et al., 2004). In a report on European survey of language testing and assessment needs of teachers, teacher trainers and experts, Hasselgreen et al. (2004) revealed that reviewing and writing tests was one of the areas which teachers and experts perceived themselves particularly in need of training.

Considering this need, the findings of this study along with the actual samples of talk might be utilized in pre-service (i.e., Language Testing and Evaluation Courses) and in-service L2 teacher education programs in the form of a testing tutorial as an awareness-raising practice and shared with student teachers or novice teachers. Those teachers can observe the IR sessions situated in authentic contexts, discuss the micro-analytic investigations of those interactions between teachers, and apply them with their mentor in their own work. In this way, they can learn the skills of IR and gain the interactional competence for testing, developing the ability to co-construct tests collaboratively. For instance, they might learn about how a test team manages disagreements as they review items and how the emerging contradictory views shape tests as well as how problems identified in the tests are dealt with interactionally. The interactional language test IR framework that is introduced in this study can also be used as a guide for effective test reviewing and for training purposes in monitoring and evaluating reviewer performance in both pre-service and in-service teacher education programs. According to Stiggins (2001), assessment practices can be taught effectively by showing real-world demonstrations of those practices in language teaching

programs and modelling them whereby student teachers get engaged in assessment preparation and communication processes. As Şallı-Çopur (2008) also suggests, since the amount of practice, experience and time is considered to be inadequate in language teacher education programs, student teachers can be presented with the research and case studies conducted in their future professional settings in Turkey to gain a sense of the atmosphere in those working environments and for the professional courses to be more relevant to the local culture and needs of student teachers. This would also help them better translate what they have learnt and practised in relation to language test construction into real testing environments where they eventually find themselves and see how teachers and testers deal with the demands of language testing in reality.

The interactional and social nature of testing that is examined in this study can be said to be in line with the new approaches in second language teacher education (SLTE) and is thus, expected to contribute to the scope of Language Testing and Evaluation Courses in those programs. Considering SLTE, the current approaches such as constructivism, transformation and sociocultural approach as well as the methodological challenges brought about by approaches such as CA have led to changes or reconceptualizations in the field and caused SLTE to be approached critically in various aspects. SLTE has come to view teachers' practices as part of the community of practice or social practice where they develop professional identities (Freeman, 2009). Therefore, there is a shift of focus to the activity of teaching within its own context and to teachers' own interpretations of their practices in the contexts where they work. This is thought to better explain the complexities of teachers' mental lives and their teaching processes (Johnson, 2009). With the reorganization in the roles of the teachers as mediators, analysts and 'independent professionals' (Leung, 2009), scholars have turned the attentions to teachers' social construction of meaning and "the character and quality of interaction between learners, teachers and the objects in their learning environment" in terms of its communicative function (Johnson, 2009, p.4). Therefore, the interactional perspective that is brought to language test construction in this study can help student teachers develop a broader testing perspective in L2 teacher education programs. They can see how language tests are co-constructed in and through talk and how such a practice is situated in interaction in an institutional context without being only limited to certain testing skills and knowledge.



### 1.3. Research Questions

Following are the main research questions of this study, which were formulated after the unmotivated examination and the initial analysis of the data. In accordance with the data-driven nature of CA research, the questions were constantly revised as the analysis continued and the interactional architecture of language test IR interactions was uncovered.

1. What is the overall structural organization of language test IR interactions in the case of problematized items?
  - 1.1. What are the phases of those language test IR interactions?
  - 1.2. How is each phase sequentially organized?
2. What testing principles do the teachers orient to throughout the phases of language test IR interactions?
3. What is the institutional role distribution of the participants displayed in the phases of language test IR interactions?

### 1.4. Terminology Used

**Language Testing/ Language Assessment:** is “concerned with procedures and techniques for measuring second language knowledge” and “while the field is generally known as ‘language testing’, it is also referred to as language assessment due to the recent trend to use multiple assessment procedures, and not just tests” (Shohamy, 1997, p.141). However, since this study is concerned with specifically language tests designed by teachers, the term “language testing” is more commonly used than language assessment. With regard to assessment as “a much wider domain than tests”, a distinction can be made between *informal* and *formal* assessment (Brown, 2001, p.402). Informal assessment “is involved in all incidental, unplanned evaluative coaching and feedback on tasks designed to elicit performance, but not for the purpose of recording results and making fixed judgments about a student’s competence” (Brown, 2001, p.402). On the other hand, formal assessments involve “exercises or experiences specifically designed to tap into a storehouse of skills and knowledge” and “are systematic, planned sampling techniques constructed to give teacher and student an appraisal of student achievement” (Brown, 2001, p.402). The language tests that this study focuses on are part of the formal assessments carried out in the English preparatory school.

**Interactional Action/ Interactional Resource:** *interactional action* is used to refer to specific actions in a turn such as suggesting, giving account, agreeing, and challenging, which become recognizable by the recipients through the way the *interactional resources* of “the language, the body, the environment of the interaction, and position *in* the interaction” are enlisted by a speaker (Schegloff, 2007, p.xiv). Considering the item reviewing practice as a project to be accomplished in Levinson’s terms, the *interactional actions* that are described are in the service of this project and thus, they are not only doing something local by means of the *interactional resources* (i.e., properties of the interaction) that govern the responses but are also part of something more global (Levinson, 2013).

**Interactional Competence:** refers to the participants’ ability to communicate in orderly and meaningful ways with one another (Drew, 2005). It is a collaborative, socially constituted domain of communicative competence including the ability to use various interactional practices such as sequencing, turn-taking and repair and to continually revise and adapt those practices in context-sensitive ways (Markee, 2000; Pekarek Doehler & Petitjean, 2017; Wong & Waring, 2010). Those practices establish the highly organized structures and patterns in conversation and thus, are necessary for the accomplishment of socially relevant acts of communication (Drew, 2005).

**Item:** “is a device for obtaining information about a test-taker’s domain of knowledge and skills or a domain of tasks that define a construct” and has two forms; “selected-response” test items where the test-taker chooses an answer (i.e., multiple-choice items), and “constructed-response” test items where the test-taker creates an answer (Haladyna & Rodriguez, 2013, p.3). Thus, in this study, it is used to refer to any question type put in the tests, whether it is a multiple-choice item, a true/false statement, or a short-answer question.

**Item Reviewing:** is the first form of item analysis and one of the stages of language test development in an educational setting where test items locally prepared for a specific content area and a group of students are carefully examined, evaluated or considered again by colleagues (i.e., teachers) with the intention of making changes if necessary.

**Problematizing a Test Item:** is the item reviewers’ act of identifying a problem requiring a solution in the language test items that are being reviewed.

**Syllabus-based Tests:** refer to tests that are prepared to assess the achievement level of the learners on the material covered in class. In those tests, the content is sampled directly from the syllabus and the test tasks represent the kinds of processes and skills that constitute the learning objectives or teaching points in the course (Fulcher, 2010).

**Test & Exam:** are used interchangeably to refer to the measurement instruments designed at the testing office to elicit a specific sample of students' language behaviour (Bachman, 1990) and to measure their language knowledge or ability in a given domain (Brown, 2001) (e.g., progress test, level achievement test, story exam). They include various test task types such as multiple-choice, fill-in-the-blanks and true/false.

**Tester:** refers to the EFL teachers who are assigned to work at the testing office of the English preparatory schools in universities in Turkey. They are responsible for preparing syllabus-based language tests at a particular level of proficiency with very few hours of teaching a week.

**Validation:** is “the process of investigating the meaningfulness and defensibility of the inferences we make about individuals based on their test performance” (McNamara, 2004, p.769).

## **CHAPTER 2**

### **REVIEW OF LITERATURE**

#### **2.0. Introduction**

In this chapter, the key concepts and issues of (1) language test construction, (2) syllabus-based language tests, (3) item reviewing, (4) interaction and teacher involvement in language test construction, and (5) conversation analysis and institutional talk as they are defined and discussed in the literature are reviewed and explained. The chapter also presents the relevant studies on interaction in language test construction as well as conversation analytic studies on teachers' collaborative workplace interactions and on language testing/assessment.

#### **2.1. Language Test Construction**

Language test construction involves the processes testers go through in modelling and eliciting the abilities that underlie learners' test performance (i.e., constructs) but that cannot directly be observed and in characterizing the essential features of performance or target language use (Douglas, 2010; Fulcher, 2010; McNamara, 2004). In this way, through the constructed test, inferences can be drawn "either about the likely quality of performance by the test-taker under non-test conditions, or about the test-taker's standing in relation to a relevant domain of knowledge and abilities" (McNamara, 2004, p.765). The fundamental goal of language test construction is, thus, to ensure that the scores or the information obtained from tests are reliable, valid, useful and fair and that the test performance is appropriate and meaningful for the intended test interpretations and uses (Bachman, 1990). "Identifying, estimating, and controlling the effects of factors that affect test scores" is the major concern in the design and development of language tests (Bachman, 1990, p.160).

The stages and models of language test development and construction have been defined clearly in the literature to provide a principled basis (e.g., Bachman & Palmer, 1996; Brown, 2001; Carr, 2011; Douglas, 2010; Downing, 2006; Fulcher, 2010; Fulcher & Davidson,

2007; Hughes, 1989; Izard, 2005; Milanovic, 2002; Mislevy et al., 2002). For instance, one of the most comprehensive models is Fulcher's (2010) test design cycle, which consists of eight components: (1) the statement of the test purpose, (2) the definition of the test criterion, (3) construct definition, (4) task/item design: specification writing, (5) evaluating, prototyping and piloting, (6) field testing, assembly into final format, (7) inferences, and (8) decisions. In those models of language test construction, at the basic level, the stages involve clarifying the purpose, defining a construct, writing the test specifications, writing the items, reviewing and revising them, piloting the test and analysing the test results, and finally producing a number of test forms and turning them into measurement instruments, which is then followed by validation studies and large-scale trialling (Buck, 2009). However, despite the suggested linear steps, language test construction, in fact, involves a thorny and an iterative process. This is made clear by Bachman and Palmer (1996) who state that

in practice, although test development is generally linear, with development progressing from one stage to the next, the process is also an iterative one, in which the decisions that are made and the activities completed at one stage may lead us to reconsider and revise decisions, and repeat activities, that have been done at another stage (p.86).

As Fulcher and Davidson (2007) also point out, language test development process is much messier than it is thought and the test is not always more reliable and valid if the tester works through the linear steps because all the activities influence each other and take place in a cyclical fashion, leaving a room for continuous change. Therefore, in language test development, the processes should not be considered as linear or fixed but as iterative and dynamic, requiring an ongoing quality control, constant checking, reflection, and negotiation (Buck, 2009; Kim et al., 2010). For instance, throughout the process, even after the completion of test specifications, some ideas for tasks may be rejected and different task types may be suggested or the construct definitions may be modified, causing the test specifications and test tasks to evolve over time (Fulcher, 2010).

The construction of a high-quality test that validly fulfils the intended functions is a complex task for test developers and involves many considerations (Linn, 2006; Welch, 2006). Items as the major building blocks for tests should be created and designed in such a way that they measure the important content at the right cognitive level and are presented in the appropriate item format (Downing, 2006). The tasks should provide clear evidence of students' knowledge and skills as stated in test specifications, make clear what is required of

the students, not be confusing and offensive and engage the students well (Izard, 2005). Since every facet of a test task influences the processes test-takers go through and may cause variance in responses, decisions about the wording of a prompt, the context, the timing and the response mode should be made with great care (Schoonen, 2011). Therefore, good item writing requires creativity, sensitivity, insight and imagination as well as experience in teaching students who are similar to test-takers (Alderson et al., 1995). It also requires a great deal of skill and knowledge about the curriculum and about how students learn a foreign language (Izard, 2005). Item construction is, thus, considered to be “more art than science” although the principles of item writing provide a scientific basis (Downing, 2006, p.10) and is viewed as a matter of problem solving (Hughes, 1989). Gronlund (1985) also sees good item construction as an art as it requires “a thorough grasp of subject matter, a clear conception of the desired learning outcomes, a psychological understanding of pupils, sound judgement, persistence, and a touch of creativity” (p.146).

What is more, when it is the language ability that is tested, test construction becomes even more challenging and complex because “language ability per se is a complicated, multifaceted construct, consisting of many interdependent or independent subskills, and each subskill can be operationalised in many different ways” (Schoonen, 2011, p.701). Similarly, a unique characteristic of language tests that could cause a dilemma for language testers is that language is both the instrument and the object of measurement, which means that the characteristics of the instrument or the elicitation procedure will overlap with the characteristics of the language abilities measured, affecting language test performance (Bachman, 1990). Therefore, due to the nature of language, it is almost impossible to write ‘pure’ items which test a single construct.

What adds to the complexity of language test construction is the impact and far reaching consequences of tests on individuals at the micro level and on society and the educational system at the macro level. This is because tests are used in contexts with specific values and goals and for various decisions and thus, act as instruments of power and control (Bachman & Palmer, 1996; McNamara, 2004; McNamara & Roever, 2006; Shohamy, 2007; Weir, 2005). In this respect, language testing is a complex social phenomenon and practice (Fulcher, 2010; McNamara & Roever, 2006). According to Bachman and Palmer (1996), at the micro-level, test-takers can be affected by the testing procedure, the feedback they get about their performance on the test and the decisions made about them based on the test

results. For example, the topical content of the test task can negatively affect the test-taker if it presents topical and cultural information that is new to the test-taker. Teachers are also directly affected by tests because tests have an impact on the program of instruction (i.e., washback effect), which might result in positive impact when there is correspondence between the characteristics of the test and those of the instructional program or negative impact when teachers start to “teach to the test”, emphasizing what is measurable and tested and narrowing the curriculum (Bachman & Palmer, 1996). At the macro-level, due to the impact of tests on society and education systems, the tests should be developed in consideration of test consequences, societal and educational values and goals and are not to be treated as isolated events (Bachman & Palmer, 1996; Fulcher, 2010; McNamara & Roever, 2006; Shohamy, 2007). In this way, the relevance and usefulness of the test scores to the decisions can be questioned and the inferences drawn from a test score about the knowledge and abilities of a test-taker can be justified based on a convincing argument (Fulcher, 2010). Therefore, language test construction is a critical process and a practice that involves careful thought.

Considering the iterative and dynamic nature, the complexity and the impact of language test construction, it is clear that it involves many considerations and a critical eye. It requires much more than testing skills, knowledge and principles and cannot simply be realized by following a prescribed step-by-step standard test development procedure and using popular language tests as models (Bachman & Palmer, 1996). As Davidson and Lynch (2002) make clear, simple recipes will not work for test construction because they are not likely to fit all settings and it is best for test developers to write their own recipes. Similarly, Buck (2009) states that “each project is different, and the good test developer is the one who can adapt to the requirements of each project, and still produce a good product despite the constraints” (p.183). Thus, creating real tests is a challenge that “requires rigorous thinking, intellectual creativity, complex problem-solving, considerable empathy, high-levels of diplomacy, a good business sense and sound project management skills” (Buck, 2009, p.184). This study also demonstrates the complexity of language test construction and the need for various skills by analyzing the item reviewing interactions that take place between EFL teachers throughout the test development process, which play an utmost importance in shaping tests within their own social context. With a micro-analytic focus on those interactions, this study is, thus, expected to present a realistic picture as to how tests are co-constructed in process.

## **2.2. Syllabus-based Language Tests**

This study focuses on the construction of in-house syllabus-based language tests. The purpose of those tests is to assess the achievement level of learners on the material covered in class. In syllabus-based approach, when planning and writing tests, the content is sampled directly from the syllabus and the test tasks represent the kinds of processes and skills that constitute the learning objectives or teaching points in the course. The consideration of the context of learning and in-class activities is highly important in order to minimize the possible negative bias of test method and maximize the positive washback of testing on teaching procedures (Bachman, 1990).

Of the syllabus-based language tests, progress achievement and final achievement tests are the focus of this study. Final achievement tests are those administered at the end of a course of study to identify how well students have mastered course content and met course objectives, whereas progress achievement tests are administered in certain time intervals throughout the course of study as they aim to measure students' achievement of a set of short-term objectives and their progress towards course objectives (Carr, 2011; Hughes, 1989; Alderson et al., 1995). Those tests are directly related to the classroom lessons, units or the whole curriculum of a specific language program as well as the actual materials that learners have been studying and aim to identify the success of the students as well as the courses in achieving the explicitly stated objectives of the program at the end of a period of instruction (Brown, 2001; Douglas, 2010; Heaton, 1990; Henning, 1987; Hughes, 1989).

Following a syllabus-content approach (Hughes, 1989), achievement tests are based on a detailed course syllabus, which is prepared in line with the course objectives and which provides the theoretical definitions of the language abilities to be tested (Bachman, 1990). Therefore, for the test to be fair, it contains what is taught to the students as part of the syllabus and that the items of the tests are drawn from the content of instruction directly. Such tests can provide both teachers and learners with information about how much learning has taken place in a certain period of time, what specific objectives have been met and whether adequate progress has been made (Douglas, 2010; Heaton, 1990). In this respect, achievement tests may serve as diagnostic tests, enabling teachers and administrators to make the necessary changes in the curriculum design, materials and the teaching methodology or practices based on the test results and to do remedial work on a certain unit



of study where necessary (Bachman, 1990; Douglas, 2010; Henning, 1987). Since the results of achievement tests can be used to both assign grades to students at the end of a unit, course or program and decide whether to review the material further before moving on, they can perform both summative and formative functions (Carr, 2011). Moreover, achievement tests are the tests that teachers are most often involved in constructing, administering, and/or scoring and their construction involves formal assessment procedures but is relatively easy since what is to be tested has been defined by the syllabus the test will be based on (Alderson et al., 1995). Thus, the test designers have to adequately sample the syllabus or the textbook in terms of content and method.

### **2.3. Item Reviewing (IR) Practice in Language Test Construction**

Most of the models of language test development include item reviewing and editing as one of the stages (e.g., Bachman & Palmer, 1996; Carr, 2011; Downing, 2006; Fulcher, 2010; Fulcher & Davidson, 2007; Hughes, 1989; Izard, 2005; Milanovic, 2002) and although different terms are used to refer to it, in this study, the term item reviewing (IR) is used. For instance, Hughes (1989) calls it moderation, Bachman and Palmer (1996) talk about try-out phase, Milanovic (2002) explains the stage of detailed editing of tests with the term vetting and Fulcher and Davidson (2007) and Fulcher (2010) use the term alpha testing or alpha prototyping for item review. Weir (2005) refers to experts' or colleagues' evaluation of tests to establish reliability of judgements as "expert scrutiny".

Item review as the first form of item analysis is an essential part of language test development process. According to Fulcher and Davidson (2007), each item in a test should be reviewed as it is an important way of checking the quality of the work and thus, is a critical part of the quality assurance process. To identify additional problems in the content and form of each item in a test and ensure that the items make sense for the purposes and the content of the course or program, the test should be subject to expert or colleague judgement as part of the item quality analysis before it is actually used (Brown, 1996). Critical item review and editing is significant because "even skilled and experienced item writers sometimes produce flawed items that are ambiguous, have no correct answer, or are unintentionally offensive to some groups of test-takers" (Linn, 2006, p.32). This is also acknowledged by Baranowski (2006) who indicates that even if an item is carefully written by experienced item writers, it should be reviewed extensively, revised and re-read well in

advance of administration. In developing tests, there is a need for additional knowledge sources since the testers' knowledge is incomplete (Shohamy, 2001). In relation to this, Alderson et al. (1995) make clear that it is not possible to produce a good test item without advice. What seems clear to a person may not necessarily be clear to others. Because the item writer knows what the item is intended to test and what the correct answer is, s/he will find it difficult to view the item objectively, not realizing that the item might in fact be testing something else and will have "quite a different view of how students will or should process the item from somebody who does not know what the 'correct' answer is" (Alderson et al., 1995, p.63). Therefore, Milanovic (2002) stresses that an independent editor or a colleague, who is familiar with the general format of similar tests but who has not been involved in the construction of the test, can provide useful feedback about the test.

Content review and professional editing of test items offers an important source of validity evidence and thus, is a necessary validity requirement (Downing, 2006). Qualitative item review by a panel of experts is one of the common methods used for providing different types of validity evidence such as content-related, construct-related and fairness-related evidence (Haladyna & Rodriguez, 2013; Wolf et al., 2008). Instead of statistical results, the step of item reviewing presents procedural results and evidence that contribute to the validation of test items in the item development process (Haladyna & Rodriguez, 2013). That is, the procedures for developing and improving test items provide a primary form of validity evidence that supports the claims about the interpretation and use of a test score and help to ensure that the item has high quality (Haladyna & Rodriguez, 2013). Therefore, the practice of item reviewing as one of the stages of test development is part of the argument and evidence for item validity as items are reviewed in terms of, for instance, their content, cognitive demand, and adherence to item writing guidelines and then are either kept as they are or revised and improved or alternatively discarded depending on the outcome of the review process (Haladyna & Rodriguez, 2013).

Baranowski (2006) states that "if item writing is considered an art, item editing is a craft, which can improve the overall validity of a test" (p.349). This is because the review of test items by a professional test editor improves the accuracy and clarity of the questions and often helps deal with issues or flaws in item content (Downing, 2006). It is also because item review minimizes the effect of the subjectivity of judgements and decisions involved in language test design and construction process as test developers subjectively select the

specific language abilities in a course syllabus to be included in a test, decide to use a certain item type and determine the level of items, time limits, and administrative procedures, which can cause bias, influencing the reliability and validity of a test and restricting the inferences made about an individual's ability (Bachman, 1990). As Haladyna and Rodriguez (2013) indicate, despite the availability of guidelines for item writing, it is a subjective process and the quality of items involves a great variability, as a result of which not all items survive the validation process, being either discarded or revised.

According to Davies and Elder (2005), validation is not just about following a set of technical procedures while developing a test because

they do not deal sufficiently with the test's use in a particular context for particular populations and scientific purposes, with the values of the various stakeholders (including test takers and end users) and with the relationship between the test's intended purpose and its ultimate impact or consequences (p.809).

The quantitative methods such as experimental approaches to validation examine only the products of the test-taking process (i.e., the test scores) but provide no information about the processes involved in test-taking, the investigation of which requires qualitative procedures (Bachman, 1990). As testing takes place in an educational and social context and plays a significant role in making decisions about individuals, validation investigations involve a complex process, which cannot be confined to collecting factual evidence to support the inferences made from the scores, and the educational and social consequences of test uses must also be considered and analysed through critical forms of inquiry as well as ethical questions (Bachman, 1990; Chapelle, 2012). Due to the multifaceted nature of validity, it is important to collect different sources of evidence from multiple perspectives to support the claims for the validity of a test score for a particular context of use and hence, to produce the most accurate picture of a student's ability (Weir, 2005).

In item reviewing, experts such as applied linguists, teachers with experience in teaching and testing, experienced testers and item writers who have not produced the actual items being edited evaluate the effectiveness of the items in terms of the relevance and usefulness of their content and format for their intended purpose and make suggestions to improve the items and the design of the test by either discarding or carefully revising the problematic items (Fulcher, 2010). Those experts are much more than a grammarian as they are expected to have knowledge of item writing principles and work with the item writer throughout the test development process, providing advice or guidance to improve the accuracy,

appearance, readability and fairness of test items (Baranowski, 2006). According to Hughes (1989), “colleagues must really try to find fault; and despite the seemingly inevitable emotional attachment that item writers develop to items that they have created, they must be open to, and ready to accept, the criticisms that are offered to them” (p.51). Therefore, Carr (2011) describes the reviewing of test items as the most emotionally sensitive part of the process and then highlights that item writers should not take personally the inevitable criticisms that their colleagues make with regard to their items and passages but instead accept them as part of an important and ongoing learning process. He further points out that the process of reviewing test items should be approached collegially whereby decisions are made by consensus without being enforced upon others. Likewise, Alderson et al. (1995) indicate that item writers should not be too defensive about their items and should be ready to accept the change or abandon the item if it is found too problematic or if it led to unexpected responses, working towards resolving the disagreements. Thus, what is desired in test writing team is good personal relations.

In relation to the task of the item review committees, they are expected to not simply read the test but to answer the questions as if they were students to replicate the experience of the test-takers as closely as possible and to be able to identify items that would cause a problem for students in terms of, for example, their wording and level of language (Alderson et al., 1995; Milanovic, 2002). This means that the members of the committee need to spend some time taking the test before the editing meeting and then meet to discuss each item in detail by reflecting upon their test-taking experience, comparing what they think the item is testing with the item writer’s claims and asking critical questions such as “is the task perfectly clear?”, “is there more than one possible correct response?” and “can candidates show the desired behaviour without having the skill supposedly being tested?” (Alderson et al., 1995; Hughes, 1989, p.51-52). During the meeting, they should be given the opportunity to explain (1) why they gave the answer they did, (2) how they went about answering the item, (3) what they think the item is testing, and whether (4) the level of the item is suitable for the target group, (5) anything significant in the syllabus is missing from the test, (6) the test method is familiar to the students, (7) the instructions are clear, (8) there are other possible answers, (9) the item is biased, (10) the context is appropriate and sufficient, and (10) whether there are any unpredicted problems and ambiguities in the wording (Alderson et al., 1995). They also need to make sure that the test covers the course material sufficiently and evenly and does not assess anything that has not been taught (Carr, 2011). Furthermore, as

Bachman and Palmer (1996) indicate, the evaluation of the usefulness of the test can be done either informally or formally by, for example, using checklists. Weir (2005), Izard (2005), Brown (2001) and Brown (1996) suggest using a checklist to facilitate the test evaluation as reviewers work through all the items and to make the comparison of the degree of agreement between experts more feasible. The checklists can contain questions such as “Is the wording appropriate to the level of education where the item will be used?”, “Are there statements in the items which are likely to offend?” (Izard, 2005, p.36) and “Does each item measure a specified objective?” (Brown, 2001, p.414).

Fulcher and Davidson (2007) and Fulcher (2010) present item review in terms of four sub-processes: content review, key check, bias or sensitivity review and editorial review. The aim of content review is to check whether the item matches the specification and is appropriate for the intended purpose. This is done to ensure item-specification congruence. Key check is carried out to make sure that the answer is correct, there are no plausible responses or alternative correct answers in, for example, multiple-choice items, and the suggested answers to any constructed response items are suitable. Bias or sensitivity review is concerned with the fairness of the test and involves considering whether an item, its wording and content contains any bias against a particular group of test-takers or causes any offence. It aims to reduce or eliminate any cultural, ethnic, religious, and gender sensitivities (Downing, 2006). Editorial review is checking for any spelling mistakes or grammatical errors, ambiguous wording and for the appropriacy of the language and formatting. It aims to correct deficiencies in the technical qualities of the items.

In a similar vein, Haladyna and Rodriguez (2013) describe item reviewing practice in terms of a series of reviews which include (1) reviewing the content of items to ensure that they are instructionally relevant to the domain of content being tested, (2) assessing the cognitive demand of items, (3) checking the adherence of items to item writing guidelines (i.e., a review of items for violation of item-writing rules that are adopted/adapted in a testing program), (4) editing and proofing test items for mechanics, grammar, clarity, and style, (5) evaluating item fairness to identify any construct-irrelevant factors such as difficult or inappropriate terminology, technical language, cultural references, and emotionally charged or offensive content, which might affect test performance, (6) examining if the items involve unnecessary language complexity as another form of construct-irrelevant variance that would cause a serious threat to the validity of items, and (7) verifying the key since the

reviewers may find out that there are no right answers, there is a right answer but it is not the one given in the answer key, there are two or more right answers or the right answer has variations in constructed-response items. According to Haladyna and Rodriguez (2013), the combination of these reviews is a collective effort to produce high-quality items. Such a multistage and multicomponent item review process is a highly interactive and complex system with each stage focusing on a different aspect or characteristic of the items and enables the review of items from diverse viewpoints, bringing unique perspectives and insights into the item development process (Welch, 2006).

## **2.4. Interaction and Teacher Involvement in Item Reviewing Practice**

Interaction with stakeholders has been highlighted in constructing and reviewing language tests and is thought to provide important qualitative evidence for item reviewing and editing. It is acknowledged that the best tests are created by a group of people through their collaborative effort rather than by individuals in isolation (Davidson & Lynch, 2002). Compared to individuals working alone, a team of item writers can construct better items and with desirable qualities because “item writing without the benefit of interaction with colleagues is generally inefficient and tends to be too idiosyncratic, representing only one person’s limited view of the topic to be assessed” (Izard, 2005, p.15). Since each person’s knowledge, experience, values, constraints, and priorities determine their choices, dialogue with professional peers will support the decision-making (Hamp-Lyons, 2000). Because everyone brings their own preconceptions, understanding and assumptions to the assessment process, the act of assessment is viewed as interpretive, involving interpretation at every point throughout the test construction process and thus, through dialogue and a mutual process of making sense of each other’s contributions, those interpretations can be uncovered and examined critically from diverse perspectives, making it possible to understand the phenomenon at hand (Scarino, 2013).

Fulcher (2010) also emphasizes the review of sample items in teams with test developers, applied linguists and teachers through critical discussion in order to investigate the usefulness and usability of the test. In this way, an agreement can be achieved as to whether the item will work and elicit a response that is appropriate for the construct. Fulcher (2010) stresses that item development and review should be considered as a collaborative group activity because it is not possible for individuals to see all the problems with items and tests

and those problems are often realized in discussion with group members, who then come up with solutions and ideas collaboratively through disagreements and negotiation. For instance, an option that is correct to one person may seem incorrect to another or vice versa due to different points of view and thus, having colleagues take the test and review the items can generate discussion over the correct answer and help spot additional problems before the test is administered (Brown, 1996). As Hughes (1989) points out, teamwork is considered to be the best way to improve items and identify those that need to be discarded because it is not always possible to produce perfect items. Bachman and Palmer (1996) also put forward that the teamwork is much more rewarding than the individual work in terms of both the process and the product because of the interaction, cooperation and the support that it provides. When the work is carried out in interactive teams, different perspectives and qualitative judgements can be incorporated and the items can be examined with a diversity of creative and careful thought, which gives the opportunity to discover mistakes and gaps in the tests before they are finalized. As Spaan (2007) states, when item editing is done by groups working together, the group dynamics lead to creativity “as the members of the group feed off each other’s ideas” (p. 282).

Interaction in reviewing test items is found necessary to also ensure ethicality and fairness, which have come into play with the recent view of language tests as “tools that are embedded in political and social contexts” (Shohamy, 2007, p.521). This means that language testers need to construct tests that are as fair as possible and be aware of their social and ethical responsibilities in their profession (McNamara, 2004). Ethicality and democratic assessment can be ensured through shared power, collaboration and the involvement of all stakeholders in the testing process and the implementation of locally-decided assessment practices (Shohamy, 2007; 2001). The testers’ responsibility is thus, to work together with groups of people such as teachers and peers and construct knowledge through dialogical, cooperative, interpretive and constructive sessions, which makes testing an art rather than a science (Shohamy, 2001). According to Fulcher and Davidson (2007), language testers are considered to be part of a community of practitioners and are expected to take part in discussion, open debate and research for the test development to be an ethical practice. As a democratic approach to testing, they need to listen to and take into account contrary views and new evidence and make sure that the test is accepted by the community of professionals and stakeholders and that everyone’s needs are addressed through free and open dialogue and ongoing negotiation (Mathew, 2004). As Fulcher and Davidson (2007)

further highlight, “the test development team and all its stakeholders should never be prohibited from discussing any aspect of testing, and provided that they keep the test effect in mind, we believe they have done their job” (p.148).

Reviewing language tests in and through interaction may also assist in identifying threats to test validity and thus, is necessary for effective test validation, which is “the process of investigating the meaningfulness and defensibility of the inferences we make about individuals based on their test performance” (McNamara, 2004, p.769). To ensure that the test is valid, it is necessary to critically evaluate, defend, challenge, and question the claims about the abilities of test-takers because of the indirect relationship between the test and the domain of knowledge or ability that we want to gather information about (McNamara, 2004). Therefore, the role of the language tester is to collect evidence to support test use and interpretation that is accepted by the stakeholders and eventually to ensure that the use of the test is appropriate for its stated purpose and target population (Fulcher & Davidson, 2007).

As Fulcher and Davidson (2007) further indicate, for test validation, language testers need to create arguments for the claims of test score meaning and the interpretations that they make about individuals, which they name as “validity argument”. More specifically, they are required to make the argument that the scores “are useful for telling us about learners’ ability to use the target language, and to do so in the context of the TLU [target language use] domain to which we wish to generalize” and that the decisions to be made with regard to the score-based inferences are appropriate (Carr, 2011, p.158). These claims need to be supported and justified by presenting evidence of the test’s authenticity and construct validity as well as test’s impact (Carr, 2011). Although there are many kinds of evidence that can be collected to support the validity argument, evidence for authenticity, for instance, can be provided through a comparison of the test tasks with the relevant TLU domain tasks; construct validity can be ensured by comparing test content with course or syllabus content in classroom based testing, and test’s impact can be evaluated in terms of its influence on the way students are studying and teachers are teaching and on the program or school in general as well as in terms of its fairness, including absence of bias and social consequences (Carr, 2011).

Such a validation process means that there will also be disagreements and alternative explanations that challenge the validity argument. Thus, test validation should involve



investigating those alternatives and successfully eliminating the counter-arguments to explain the evidence and support the use of a test for a given purpose and the interpretation of scores (Fulcher & Davidson, 2007; So, 2014). In other words, the contradictory evidence that arises in validation research should be removed by redesigning or revising the test (Davies & Elder, 2005). In doing this, testers' interaction with stakeholders is important and an argument is said to support the claims about the meaning and use of a test "if it is acceptable to the community of language testers and stakeholders in open discussion, through a process of dialogue and disagreement" (Fulcher & Davidson, 2007, p.21). Likewise, Chapelle (2012) indicates that the demonstration of the validity of test inferences and uses takes place in a social context where professionals together agree on the acceptable procedures.

Teachers' involvement in particular in the IR practice and interaction with them has been considered as an important part of language test construction. Firstly, teachers know a great deal about the abilities, knowledge and needs of their learners as they spend a lot of time interacting with their students and observing them participate in various activities and tasks in the classroom (Fulcher & Davidson, 2007). Since they are the ones who have taught the material, they can easily help devise test tasks that reflect what learners have been taught and how the material has been presented and practised in class, which is especially important in the design of syllabus-based classroom language tests. As they develop insights about learner progress with experience, teachers often become skilled judges in evaluating language performances and in making diagnoses about learners' difficulties, and thus, they can be effective in evaluating the quality and usefulness of the tests in their institution (Brindley, 2001). In this respect, teachers can be considered to be the best sources of information about the students and the effectiveness of teaching and learning. Item writers, who work full-time for testing units, may be less sensitive to the test population, to the changes in the curriculum and its implementation and to other features of testing and teaching context and thus, the best solution is to have item writing teams that involve experienced teachers (Alderson et al., 1995). Therefore, including teachers in the test development process can help identify any mismatches between teaching and testing (Lynch & Davidson, 1994). The design of tests with a consideration of the learning context, student cohort and subject matter through teacher involvement in the process of constructing tests will, in turn, lead to better test validity, an increase in test quality, and positive washback effect on instructional practices (Dikilitaş et al., 2018; Xerri & Vella Briffa, 2018).

Secondly, one advantage of involving teachers who teach the course for which the test is designed in IR process is that it will enable teachers to feel more responsible and to be more cooperative in the testing process as they think about the link between what is taught and what is tested and as conflicting views arise and they try to achieve compromise (Brown, 1996). In relation to this, Davison (2004) points out that teacher interaction on assessment issues is a good opportunity for teachers “to develop a sense of ownership and common understanding of the assessment process and to articulate and critique their own implicit constructs and interpretations”, and as a result, to have a more informed perspective of assessment practices through interaction, dialogue and trust and to rely on tests as accurate measures of learners’ progress (p.328). As Xerri and Vella Briffa (2018) indicate, teachers’ involvement in testing is not only facilitated by their assessment literacy but also contributes to the development of their assessment literacy, which has a direct link to positive student outcomes since their understanding of the testing process has a positive impact on their classroom practices and their attitudes to assessment.

Thirdly, teacher involvement in the IR practice allows teachers’ voices to be heard in a prescriptively oriented educational system, giving them the chance and power to influence language testing and to keep a track of their learners’ progress effectively. Teachers want assessments that can “provide sound judgements of learning based on an understanding of learners’ needs and background” (Hamp-Lyons, 2007, p.493). For such assessments to take place, there needs to be “high contextual awareness and responsiveness”, and an involvement of those who are actually engaged in teaching in assessment (Hamp-Lyons, 2007, p.493). As Shohamy (2007) also indicates, as a democratic approach to assessment, language teachers’ participation and representation in the decision-making process in testing is highly important. In this way, multiple interpretations can be incorporated and the power can be shared, allowing the teachers to become skilled and confident in testing. Teachers’ involvement in the testing process will, thus, empower them and encourage confidence and creativity in their own judgement (Xerri & Vella Briffa, 2018). This is also put forward by Fulcher (2010), who states that “teachers need to become strong advocates for change and for social justice, rather than bystanders to whom testing ‘happens’” (p.8). Finally, testing within local contexts is required to be “ecologically sensitive”, which means that it should fulfil the local needs of teachers and learners and make the interpretations of the meanings of test scores in relation to a particular learning environment (Fulcher, 2010). One characteristic of such “ecologically sensitive” tests is the involvement of local stakeholders

such as teachers in constructing and implementing tests and the correction of mistakes in tests through dialogue and negotiation (Fulcher, 2010).

The practice of reviewing language tests, the role of interaction, collaboration, and teamwork in doing this and teachers' involvement in IR interactions especially in the case of syllabus-based language test development have been acknowledged and discussed as an important part of language test construction process. The IR practice has generally been explained in terms of sub-processes and checklists have been suggested to guide those processes. However, the micro-details of how actually IR interactions are realized, how language tests are conjointly constructed in and through interaction with classroom teachers in a specific institutional context and how this shapes the design of language tests are unknown. This study aims to contribute to this gap and shed a light on the interactional processes of IR practice in language test construction, bringing to the fore the role of professional interaction between classroom teachers and testers. In doing so, this study uses Conversation Analysis as a methodology to analyze this interactional nature.

## **2.5. Studies on Interaction in Language Test Construction**

Language assessment research has come to utilize various research methods with a shift of focus from positivistic research and traditional statistical methods to a more critical examination of a range of validity issues (Lumley & Brown, 2005). For instance, a variety of qualitative approaches have been used to investigate a range of test-related issues by consulting with stakeholders, analysing not only the product but also the process of language testing and paying greater attention to the consequences of language assessment (Lazaraton & Taylor, 2007; Lumley & Brown, 2005; McNamara, 2006). Besides psychometric methods, those qualitative methods are necessary to gain a fuller understanding of language tests (Lazaraton & Taylor, 2007). In language testing studies, the qualitative research methods that have been introduced are discourse analysis, conversation analysis, observation checklists, verbal protocol analysis and interviews, which have been mainly used to shed light on interlocutor behaviour during the process of speaking assessment, test-taker behaviour, and on rater behaviour in the process of writing assessment (Lazaraton & Taylor, 2007).

Regarding the study of stakeholder feedback and involvement in language test development, so far language test construction and validation studies following a qualitative and process-oriented approach have investigated both written (e.g., surveys, written comments and reports) and oral (e.g., interviews, focus groups, group discussions, meetings, workshops) feedback primarily from test-takers (e.g., Erickson & Aberg-Bengtsson, 2012; Hamid, 2015; Mathew, 2004; Ockey et al., 2013; Suryaningsih, 2014) and from teachers as in this study (e.g., Erickson & Aberg-Bengtsson, 2012; Green & Andrade, 2010; Lynch & Davidson, 1994; Mathew, 2004; So, 2014; Tan & Turner, 2015). Since the focus of this study is on interactions that take place between teachers and item writers at the stage of IR in language test construction process, the studies that are specifically concerned with the form of feedback obtained from colleagues in and through interaction are considered relevant and described below. Those studies dealing with teacher or any other kind of expert interactive feedback in language test construction used a case study methodology particularly through workshops, focus groups, meetings, and small group discussions and mainly focused on large-scale tests. They examined how the incorporation of English teachers' perspectives, collaboration and feedback and the interactional processes involved in doing this in terms of joint decision-making and effective communication influenced the development and validation of those tests. However, they did not have a special focus on IR practice and on the sequential unfolding of teachers' interactions.

In a case study, So (2014) examined how the incorporation of English teachers' perspectives into the development of a large-scale English test through focus group interviews at the stage of piloting the trial items and how their feedback can support the validity argument for score interpretation and test use. According to So (2014), although scholars in language testing emphasize the need for stakeholders' perspectives at various stages of test development, there is little research that examined whether and how stakeholder feedback was obtained and applied in practice in language test construction process. The study showed that the teachers, who were the stakeholders involved in focus interviews, acted as critics and their feedback involved suggestions for revision and improvement of trial items in ways that would make the items appropriate for the target population and purpose, providing useful evidence for the validity argument and making significant contributions to test development. So (2014) indicates that stakeholder feedback is important because different groups of people may have different opinions and concerns about the use of a test and those opinions might be in conflict with those of others, influencing the acceptability of

the test (So, 2014). Thus, in the study, it is pointed out that gathering stakeholder perspectives about items can be an effective way of identifying any potential threats to the validity of test score interpretation and use. Of the stakeholders, this study particularly gave importance to teachers' perspectives because they "have the best understanding of the potential test takers and the English language curriculum and materials", and are able to verbalize learners' concerns, becoming their voices (So, 2014, p. 292).

Via a workshop approach, Lynch and Davidson (1994) focused on teacher feedback and its contribution to the consequent refinement of test specifications and items in criterion-referenced language test development, giving priority to teachers' knowledge and experience in collecting preliminary evidence for the test validity. They found that during the workshop experiences, the teachers had the chance to verbalize the gap between their teaching methodology and the test by discussing the accuracy versus fluency issue, to uncover mismatches between the metalanguage used in the textbook materials and the test and decide on the terminology to be used in the test and to clarify their instructional objectives in an attempt to match the test with the goals of the institution. According to Lynch and Davidson (1994), such a process in criterion-referenced language test development leads to the empowerment of teachers.

In another study, Davidson and Lynch (2002) investigated the group dynamics of testcrafting and more specifically, how group dynamic variables were manifested in a language test specification development by drawing on discussions of group structure and group process in the small group research literature. The test specifications were drafted and finalized in two groups of nine people in an eight-hour workshop, which was videotaped, and then the groups participated in a debriefing session to discuss their reflections on the experience and the future use of the test specifications. The participants were also interviewed individually about their perceptions of the experience. The findings revealed that (1) the participants did not assume certain roles and their contributions to the discussion changed depending on their teaching, testing and language learning experience and on the need for such expertise without any one of them dominating the group's interaction, (2) the participants made an effort to create and maintain group cohesiveness; (3) one of the groups showed more group maintenance activity in an attempt to respond to one participant's questioning behaviour, (4) although the groups managed to critically evaluate the consequences, certain characteristics of their solutions and the biases of their solutions (e.g.,

how the test specifications would affect item writers, teachers and students), they did not consider and critically examine a range of alternative solutions and the major values for the choices they made, and failed to develop contingency plans, which were likely to be due to time constraints, (5) the groups used certain interaction and communication strategies to deal with the ambiguities and obstacles, and (6) of those strategies, the groups used “Devil’s Advocate” approach in questioning each other’s ideas and checking the effectiveness of the explanations offered.

According to Davidson and Lynch (2002), there is a need to raise awareness about the characteristics and strategies for managing and promoting group cohesiveness in the contexts of testcrafting. In this way, participants can be presented with the effective stages as well as the types of interaction for productive group development and decision-making process. Participants can also be provided with the types of obstacles (e.g., disagreement, request for clarification) that are encountered in testcraft situations and with the ways of overcoming them in groups. For instance, based on small group research literature, Davidson and Lynch (2002) note that for an effective group interaction to take place, it is necessary to have a realistic range of alternative solutions and that the occurrence of disagreements within the group or the use of the strategy of playing “Devil’s Advocate” will “lead to a more extended discussion of the alternative solutions” and thus, “the generation of a greater number of ideas and solutions and to ones that may not have been thought of in a less ambiguous task context” (p.106-107). Such ambiguities and obstacles can be presented as increasing the quality of decision-making in testcrafting. On the other hand, it was shown that there can be disruptions such as when the status of a group member is used as a rationale for a decision without further discussion on the solution or when the group members behave defensively, which would be overcome by, for example, a group member reminding others of the group norms.

Using a case study methodology, Green and Andrade (2010) implemented a twelve-step model for test reform with a focus on teacher collaboration in developing end-of-semester exams in a local ESL context at tertiary level rather than large scale proficiency tests, which have been the focus of research on language testing. During the implementation of the test reform model, data were obtained through interviews, small group discussions, formal meetings, participant observation, program documents, direct observation of the test reform process and the interactions. The teachers provided feedback and were involved in the test

development process at each stage of test reform as they wrote test specifications, created, reviewed, piloted and administered tests. A leader was appointed for each group to assist the collaboration process and to make the communication between members more effective. The starting point for the study was the belief that stakeholder participation and learning is an important requirement for any type of reform and that teachers need to be included in the process to be able to accept change. According to Green and Andrade (2010), although the collaboration process in test construction can be challenging, it not only creates effective tests but also leads to “personal learning, improved program practices, more equitable distribution of power, and positive reactions to test instruments” as well as confidence in using research procedures and a sense of ownership by teachers towards the test and test development (p. 328). It gives teachers the opportunity to recognize the difficulties faced in test development and create a final product that they all accept and believe in. In this study, the test reform process was found to reveal the need to revise course objectives. It was shown that despite the conflicts that arose occasionally, it provided a valuable experience for all the staff in the program by addressing the challenges regarding assessment in their institution and designing an assessment system that is directly linked to instruction.

Mathew (2004) investigated how stakeholder (i.e., teachers and students) involvement in the design of language tests in two testing situations in India improved ethicality. The stakeholders’ views on the evaluation framework were obtained through meetings, orientation programs, review workshops, and seminars in an attempt to democratize the testing process. The study showed that the teachers and students came to terms with different aspects of the test such as test format, test difficulty and scoring criteria through open dialogue.

Tan and Turner (2015) examined the consequences of teacher involvement and collaboration in the process of the development and validation of a high-stakes ESL exit writing exam in Quebec through a mixed-methods design, focus groups and workshops with teachers at various stages of test development. The findings showed that the increased collaboration between ESL evaluation officials and classroom teachers and the teacher comments on scoring rubric and task revisions helped align the external exam and classroom-based assessment and improved the quality of the task and rubric, enhancing the exam construct validity. It was found that through workshops and rater training, teachers were able to better understand each rubric criterion and grade level as well as the demands of the writing exam.

Such a cooperative process enabled teachers to feel empowered and become more confident in guiding their students. According to Tan and Turner (2015), in the case of most large-scale test development, test-takers, teachers and the classroom context are not taken into consideration and are regarded as a source of construct irrelevant variance. However, the study highlights that as supported by many researchers, there is a need to align teaching, classroom-based assessment and external testing to improve student learning, promote teaching and to avoid negative washback because the lack of correspondence between the classroom and high-stakes contexts causes teachers to narrow the curriculum to what is tested. The study also indicates that especially in cases where the exam is developed by an outside board but is administered and scored by classroom teachers in local contexts, teachers need to be well-informed about the task requirements and assessment criteria of external tests in order to rate student work accurately.

Recent studies on language test construction also include those that focused on the collaborative processes of item writing and the effect of the item writers' interactions. Kim et al. (2010), for instance, investigated an item writing process in a case study by demystifying how item writers used test specifications, how group dynamics in a test development team affected the process, and how individual item writers' characteristics influenced item writing processes, with implications for the design of test specifications and training item writers. Since they defined group dynamics as "the interactions that take place when a group of people works for a common purpose or product", a part of the study dealt with the effect of interactions that took place in a test development team, showing how members exchanged ideas, provided feedback and made decisions in completing their shared product (p.161). The study adopted a case study approach and investigated the construction of three sets of reading tests for an online English proficiency test in two item writing groups over one semester, at the end of which the members of the project wrote a journal reflecting on their item writing experience. In this study, Kim et al. (2010) found that item writers used their own internal guidelines for item writing without sticking much to test specifications, drawing on their teaching and testing experiences, other item writing practices, face-to-face meetings, as well as oral communications with item coordinators. Thus, they showed that test specifications need to be audience-specific and designed by involving item writers in the editing and revising sessions and communicating with them for them to get a better understanding of the items to be constructed and developing "organic guidelines" that reflect item writers' trial-and-error processes. The item writing process itself also took place



collaboratively as coordinators had review meetings with item writers. In addition, it was found that the item writers' personalities, preferences, previous experience and language background influenced the feedback process and group dynamics in item writing and thus, it was suggested that those characteristics are considered in forming the groups and training item writers. Overall, the study showed that the test development is "both a consensus-building process within a team and an individual creative process" (p.161).

Similarly, through a case study, Green and Hawkey (2011) examined the processes item writers went through in selecting and adapting texts for a high-stakes test (i.e., IELTS), first in a reflective writing and a joint discussion and then in a simulated item writers' editing meeting to revise the reading texts and items. A textual analysis of the original test material and the adapted version was also carried out to be able to compare the two texts and identify the changes. As Green and Hawkey (2011) pinpoint, the processes that item writers go through in generating tests are under-researched and they aimed to make a contribution to this area in language testing. It was found that in selecting a text, the item writers all made use of item writer guidelines document and considered both the nature of the content and its technical suitability in terms of the level of language and the density of the text. Following this, the editing meeting demonstrated the complexity of the decision-making processes in preparing reading tests and resulted in editing actions such as reducing redundancy and technical language, changing styles, deciding on sensitive issues and relationships between texts and test items in an attempt to make the texts more academic and thus, relevant for the target language use domain and to include sufficient distraction for selected-response items. The observation of the item writers in the simulated editing meeting showed how the group cooperated and what features of texts and items they criticized, rejected or modified. In this study, it is emphasized that item writers need feedback from colleagues to be able to successfully meet the requirements of the institution and that expertise itself is not sufficient to ensure test quality. It is also noted that the editing meetings provide opportunities for interaction and contribute to professional development.

To summarize, those studies have a number of outcomes that are relevant to this study with respect to the role of interaction in constructing tests. They show that receiving feedback from colleagues in and through interaction during language test construction process (1) provides useful preliminary evidence for validity argument and is an effective way of identifying any potential threats to the validity of test score interpretation and use by

uncovering any mismatches between teaching and testing and addressing the challenges regarding tests in their institutions (Green & Andrade, 2010; Lynch & Davidson, 1994; So, 2014; Tan & Turner, 2015), (2) helps to criticize, reject or modify certain features of texts or items and to eventually come to terms with different aspects of tests such as test difficulty and scoring through cooperation and open dialogue (Green & Hawkey, 2011; Mathew, 2004), and (3) leads to a productive decision-making process as the participants use certain interaction and communication strategies (e.g., playing “Devil’s Advocate”) to create and maintain group cohesiveness, to generate an extended discussion, and to deal with the ambiguities and obstacles (e.g., disagreement) that arise in the group (Davidson & Lynch, 2002).

Although those studies dealt with teacher feedback and the role of interaction in language test development as in this study, they did not specifically investigate the stage of IR except So’s (2014) and Green and Hawkey’s (2011) studies but rather the design of language tests as a whole (Green & Andrade, 2010; Mathew, 2004; Tan & Turner, 2015) and such stages as determining test specifications (Davidson & Lynch, 2002; Lynch & Davidson, 1994) and item writing (Green & Hawkey, 2011; Kim et al., 2010). They also mainly focused on large-scale tests, and used primarily case study methodology through, for example, focus groups and interviews. As a result, to the best knowledge of the researcher, the micro-details of the actual interactions that take place between teachers and testers in the process of developing language tests in local contexts have not been scrutinized. In other words, the interactional processes item writers go through in generating an in-house language test remain under-researched. Therefore, this study aims to make a contribution to the studies on interaction in language test construction by carrying out a micro-analytic investigation of the IR interactions between teachers and testers via CA. In this way, it is believed that the processes of IR can be better explained, providing a more vivid picture of syllabus-based language test construction.

## **2.6. Conversation Analysis (CA)**

Conversation analysis (CA) has opened a new way to the study of interaction. Since we accomplish our social lives through talk-in-interaction and “life’s experiences- the ordinary and the everyday, the profound and the momentous- are first and foremost experiences that are shared socially through the activity of talk” (Gardner, 2004, p.263), the issue of

interaction has received a deal of attention as an object of inquiry in its own right over a long period of time and has been examined from various perspectives such as Gumperz's Interactional Sociolinguistics and Hymes's Speaking Model. However, as Liddicoat (2007) argues, most of the approaches to conversation offered a set of prescriptive social rules and principles describing how a conversation should be and how language should be used and thus, "much of what has been written about conversation is prescriptive in nature" (p.1). Liddicoat (2007) further states that they "frame conversation as an elite activity governed by the conventions of 'polite society'" while it is true that conversation is an everyday activity and a structured social event (p.1).

The roots of CA are traced back to Sociology in the 1960s with the work of Sacks (1992) and his collaborators Schegloff and Jefferson, who were the leading conversation analysts. However, since then, CA has earned an interdisciplinary reputation and has been applied in a variety of disciplines and fields such as applied linguistics, anthropology, discourse analysis, pragmatics, sociolinguistics, psychology, and communication studies (Psathas, 1995). Sacks's view of sociology and thus, CA was highly influenced by Goffman's interaction order of face-to-face communication and Garfinkel's ethnomethodology as a field of Sociology (Hutchby & Wooffitt, 1998; Liddicoat, 2007; Ten Have, 2007). CA drew up and grew out of developments in these two fields, which also had a concern for scrutinizing the orderliness of everyday life (Psathas, 1995; Liddicoat, 2007). Based on his observations of people in interaction, Goffman, for example, showed that there was a great social significance in ordinary activities of daily life such as presenting oneself to others, face-to-face gatherings, and the organization of traffic on public sidewalks (Psathas, 1995). Similarly, Garfinkel's focus on practical reasoning in everyday activities and his interest in procedural study of common-sense practices, which reflected a series of patterns but which were not necessarily interactional, by means of field observation, coding and categorization later led the way to CA upon the discovery of the idea that the action an utterance performs depends on its sequential position (Liddicoat, 2007; Ten Have, 2007). It was realized that talk was itself action based on the observation that participants could achieve a coherent, rational and mutually comprehensible interaction (Gardner, 2004). Moreover, although CA has derived from ethnomethodology, they have different research agendas (Markee, 2007). While ethnography "ultimately seeks to develop grounded understanding of *why* members act in particular ways", CA "strives to provide sequential analyses of *how* participants achieve particular practices" (Markee, 2007, p.1023) since the orderliness of interaction is

available to both participants and analysts as an achieved orderliness, providing a methodological resource to examine participants' accomplishment of social actions through the interplay of the body, talk and material resources (Heath & Luff, 2013).

As Wong and Waring (2010) explain, the main characteristic of CA is gaining an understanding of how conversation works and how participants produce joint achievements by examining minute details of recorded language and social interactions and how members orient to the practices of the sequential organization of talk, turn taking, and repair, which underlie the co-construction of talk-in-interaction (Markee, 2007). According to Walsh (2011), the aim of CA is "to account for the structural organisation of the interaction as determined by the participants" without making an attempt to analyze the data based on pre-determined categories or external concepts such as intentions, beliefs and emotions that cannot be shown in the participants' observable behaviour in the data (p.88). Therefore, CA offers an emic perspective in the sense that the perspective is embedded in the spoken interaction itself and is revealed based on participants' interactional practices, organizations and orientations rather than their perspectives obtained through, for example, interviews or the researcher's perspective (Markee, 2007; Seedhouse, 2004, 2005; Ten Have, 2007; Walsh, 2011). In other words, what makes CA emically oriented is its access to the participants' own understanding and interpretation of what they are doing through a micro-analytic description of their turns-within-sequences without making any macro-social assumptions unless otherwise displayed or made relevant by the participants in and through talk (Gardner, 2004). That is, CA studies the very same features of conversation that speakers use in producing and understanding it and thus, what is available to speakers through their turns-at-talk and their displayed understandings of each other's talk on a turn-by-turn basis is also available to the analysts, which provides an empirical foundation for the study of talk-in-interaction (Sacks, 1992). Hutchby and Wooffitt (1998) provide the following explanation for CA:

Conversation analysis is characterized by the view that how talk is produced and how the meanings of talk are determined are the practical, social and interactional accomplishments of members of a culture...participants in conversation are seen as mutually orienting to, and collaborating in order to achieve orderly and meaningful communication. The aim of CA is thus to reveal the tacit, organized reasoning procedures which inform the production of naturally occurring talk (p.1).

More specifically, Hutchby and Wooffitt (1998) define CA as the systematic and fine-grained analysis of recorded, naturally occurring talk produced in everyday situations of

human interaction across different contexts in order to “discover how participants understand and respond to one another in their turns at talk, with a central focus being on how sequences of actions are generated” (p.14). Thus, the main object of study in CA is the interactional organization of social activities and the endogenous analysis of patterns in this organization reveals how people manage and accomplish the sequential order of talk-in-interaction. In Psathas’s (1995) terms, social actions have a natural or intrinsic organization and CA is interested in finding “the structures that produce and constitute that orderliness” (p.2) through a moment-by-moment analysis of talk-in-interaction in terms of sequences, turn-taking and repair practices (Markee, 2000). In brief, CA is “the study of sequences of action and their interactional products” so that it can be revealed “what it is that makes some talk just what it is, and nothing else” and what underlying structures the members of a society draw on to construct their socially distributed activities and to engage in mutually intelligible social interaction (Gardner, 2004, p.266). As Drew and Heritage (1992) point out, CA focuses on “the interactional accomplishment of particular social activities”, which are embedded in sequences of specific social actions (p.17).

Order and structure are important issues discussed in CA, which is concerned with verbal interaction as instances of the situated social order. This is evident in Sack’s (1992) statement that there is “order at all points” in interaction (p.483), which implies that “talk in interaction is systematically organized, deeply ordered, and methodic” (Seedhouse, 2004, p.14). CA views structure as a feature of situated social interaction that participants actively orient to and they display their understanding of that order in designing their actions (Hutchby & Wooffitt, 1998). Participants treat the orderly course of interaction as normatively implemented, so departures from regulative interactional patterns and format are accountable and sanctionable (Heritage & Clayman, 2010). Structure is also defined as a local and contingent matter that is accomplished and shaped in and through the moment-to-moment sequential procedures of everyday talk (Zimmerman & Boden, 1991). Hence, with CA, Sacks aimed to reveal the levels of social order present in everyday practices of talking. He had the assumption that as a rule-governed everyday activity, ordinary conversation might be a deeply and sequentially ordered and structurally organized phenomenon rather than a random one (Hutchby & Wooffitt, 1998). According to Psathas (1995), the issue of order makes up CA’s basic assumptions as given below.

1. Order is a produced orderliness.
2. Order is produced by the parties in situ; that is, it is situated and occasioned.
3. The parties orient to that order themselves; that is, this order is not an analyst’s

conception, not the result of the use of some preformed or preformulated theoretical conceptions concerning what action should/must/ought to be, or based on generalizing or summarizing statements about what action generally/frequently/often is.

4. Order is repeatable and recurrent.
5. The discovery, description, and analysis of that produced orderliness is the task of the analyst.
6. Issues of how frequently, how widely, or how often particular phenomena occur are to be set aside in the interest of discovering, describing, and analyzing the structures, the machinery, the organized practices, the formal procedures, the ways in which order is produced.
7. Structures of social action, once so discerned, can be described and analyzed in formal, that is, structural, organizational, logical, atopically contentless, consistent, and abstract terms (p.2-3).

Another key idea in CA is the recipient design. As Liddicoat (2007) informs, recipient design takes place when the participants construct a talk in a way that is understood by the interlocutors based on their shared knowledge and common ground. Therefore, interactional procedures are designed in a way that is appropriate for recipients and CA “seeks to understand these shared procedures which participants in an interaction use to produce and recognize meaningful action” (Liddicoat, 2007, p.7). In this respect, different from the other approaches to discourse, which mainly focus on the speaker, CA views talk as a jointly accomplished activity by both the speaker and the listener as co-constructors of talk (Gardner, 2004). This is in close relation to the issue of intersubjectivity, which is considered to be an important feature of CA. The question that CA attempts to answer is; how do we share a common understanding of the world and of another’s actions in the world? (Hutchby & Wooffitt, 1998). As Seedhouse (2004) states, one of the principal aims of CA is “to trace the development of intersubjectivity in an action sequence”, which means that CA practitioners aim to show “how participants analyze and interpret each other’s actions and develop a shared understanding of the progress of the interaction” (p.13). Furthermore, Liddicoat (2007) showed that the participants might not always share the same social goals or have the same understanding of context and CA attempts to describe the progress of those participants’ intersubjectivity too.

With regard to the issue of context in CA, Walsh (2011) explains that, “social contexts are not static but are constantly being formed by the participants through their use of language and the ways in which turn-taking, openings and closures, sequencing of acts, and so on are locally managed” (p.84). He further indicates that interaction is examined in relation to meaning and context of the actions in the talk and that sequential organization of interaction

is crucial to the process of dynamically and mutually creating social context. According to Drew and Heritage (1992), from a CA perspective, “context is treated as both the project and product of the participants’ own actions and therefore as inherently locally produced and transformable at any moment” (p.19). Consequently, CA views contributions to interaction or organizational structures as context-shaped and context-renewing (Heritage, 1984a). This means that what has just been said in the preceding turn or activity (i.e., immediate local context of utterances) determines the interpretation of what is said in the unfolding talk (i.e., context-shaped), which will in turn provide the immediate context for the interpretation of the next utterance, projecting the succeeding action(s) (i.e., context-renewing) (Drew & Heritage, 1992; Gardner, 2004; Heritage & Clayman, 2010). To put it briefly, each turn is shaped by the previous turn and is itself treated as the context for the next turn. In this respect, as Drew and Heritage (1992) indicate, “the interactional context is continually being developed with each successive action” through a moment-by-moment progression of the interaction (p.18) and thus, it is the sequences of actions that heavily shape meaning and context in interaction (Heritage, 1997). As Markee (2007) states, the aim of CA is to “develop a grounded interpretation of participants’ culturally and contextually situated behaviours” by providing sequential analyses of how they co-construct talk and achieve particular practices (p.1021). Furthermore, CA is concerned with the social and contextual information (e.g., ethnographic information about the setting or the participants) that is only displayed and demonstrated by the participants as being relevant in and through talk rather than assumed (Sidnell, 2013; Zimmerman & Boden, 1991). This is most clearly exemplified by Sidnell (2013) as follows;

while it may be obvious to us as analysts and to the participants that some bit of interaction involves a ‘woman’ talking to a ‘man’ or a ‘doctor’ talking to a ‘nurse’ (etc.), it is nevertheless also true that just because a woman is speaking, she may not be speaking ‘as a woman’ (but rather as a doctor or, for that matter, as a vegetarian, a mother, an animal rights activist, etc.). As such, if the analyst wants to claim that any such identity is relevant to how some bit of talk comes off, the analyst must show that the participants themselves oriented to that categorization and designed their talk in relation to it (p.86).

Therefore, context is characterized in terms of the participants’ orientations and their documentation of it in the details of the interaction rather than a description of the physical setting or participants and the analysts cannot decide etically which aspects of context they think are relevant especially considering the fact that there are an infinite number of contextual details that could be invoked as potentially relevant (Seedhouse, 2004).

Regarding the mechanics of talk that CA aims to describe, the primary units of analysis involve turn-taking organization, sequence organization, repair and turn design, which work together to build an architecture of intersubjectivity in an interaction (Seedhouse, 2004) and to practically accomplish structure-in-action (Zimmerman & Boden, 1991). Participants use these interlocking interactional organizations “normatively and reflexively both as an action template for the production of their social actions and as a point of reference for the interpretation of their actions” in order to achieve mutual understanding (Seedhouse, 2004, p.17). Thus, they enable CA researchers to discover the underlying structures that constitute social actions; in other words, the relevant and recognizable practices participants use in the sequential organization of talk to perform the actions of everyday life. As Drew (2005) states, “the aim of research in CA is to discover and explicate the practices through which interactants produce and understand conduct in conversation” (p.75).

The practices of talk, which are a basic means of analysis and the buildings blocks of an interaction, are both context-free and context-sensitive in the sense that they are applicable to any conversation and thus, can be considered to be generic and universal but are at the same time employed locally in context-sensitive ways as the talk unfolds in a conversation by displaying an orientation to particular others (i.e., recipient design) and circumstances (Maynard, 2013; Seedhouse, 2004, 2005; Zimmerman & Boden, 1991). As a result, the context-sensitive social actions are interpreted and understood by reference to the underlying context-free machinery, which includes a set of normative resources that the analysts focus on and that the interactants use “to display the meaning of their social actions to their partners and to interpret their partners’ actions” (Seedhouse, 2004, p.38). The mechanics of talk, hence, guide the analysis and help explain how a specific event is “achieved in its particularity- in *those* words or physical actions, by *that* participant, at *that* point in the interaction, understood in *that* way by co-participants, produced by some *specifiable practices* of conduct” (Schegloff, 2007, p.252).

Central to CA is the study of turns and turn-taking as one of the interactional organizations, which refers to the basic rules of how participants take and allocate turns in interaction so that social actions can be accomplished and essential parts of meaning-making can be constituted. Turns are connected with one another in systematically and normatively organized patterns or sequences of turns and hence, turn-taking involves an orderly, rule-governed process (Gardner, 2004; Liddicoat, 2007). As Hayashi (2013) states, “conversation



is a turn-organized activity, and there are orderly ways in which participants effect transition from the current speaker to a next” (p.168). Turn organization of talk, thus, provides evidence for how participants orient to the features of a social action (Sacks et al., 1974). The rules for turn-taking organization were described by Sacks et al. (1974), which account for not only the orderliness in the turn-taking system but also breakdowns in it (Gardner, 2004). Based on the basic principle that one party speaks at a time, Sacks et al. (1974) explain that upon launching a turn, a speaker has primary rights to the floor and the transfer of speakership becomes a salient possibility at certain points. The turn-taking organization they propose focuses on two interrelated components; turn constructional component and turn allocation component. In other words, it describes how the methods of turn construction determine the points where transition to a next speaker becomes relevant and how turns are allocated, that is, the methods of selecting the next speaker.

To begin with the turn constructional component, turns are composed of the building blocks of turn-constructional units (TCUs) such as lexical items, phrases, clauses and sentences which are recognizable as possibly complete within their context. TCUs are context-sensitive and cannot be structurally defined as participants can use a range of different structures and resources depending on the context (Liddicoat, 2007; Seedhouse, 2004; Ten Have, 2007). The resources that shape TCUs and make them recognizable include grammatical features, the phonetic realization of the talk, and the action that is intended to be performed (e.g., asking a question) within the specific context (Schegloff, 2007). In addition, TCU endings, action trajectories as well as a variety of design features such as syntactic, prosodic, pragmatic and nonverbal behaviour in a TCU project the possible completion point of a turn and establish a transition-relevance place (TRP) where transition to a next speaker becomes possibly relevant (Sacks et al., 1974; Schegloff, 2007).

As for the methods for turn allocation, they include current speaker selects next (e.g., implementing an action that requires a responsive action next and the use of some form of addressing) and self-selection techniques. Those procedures described by Sacks et al. (1974) minimize both gaps and overlaps and at the same time offers a resolution of these problems so that one party speaks at a time principle can be preserved and the participants can orient to it as the norm (Liddicoat, 2007; Seedhouse, 2004). Additionally, despite the existence of some rules and principles that account for turn-taking and that provide a resource for the participants to orient to in creating their talk, the action of turn-taking is not imposed

externally on a conversation but is a social phenomenon that is locally organized and interactionally managed (Liddicoat, 2007; Sidnell, 2010; Ten Have, 2007).

On the other hand, some forms of participation in conversation are not always designed for one participant speaking at a time but involve collaborative completions and this principle is not treated as being violated on such occasions (Lerner, 2002). That is, turns can also be shared when a speaker of a turn invites another co-participant to speak in his/her turn-space through a conditional entry (Hayashi, 2013). Anticipatory completion (Lerner, 1996a; 1996b, 2004; Sert, 2019; Szczepek, 2000; Van Kruiningen, 2013) and choral co-production (Lerner, 2002) are two of the interactional practices that demonstrate a conditional entry into a speaker's turn-space. Anticipatory completion is a practice where the recipient of an ongoing turn that is produced in a compound turn-constructive form (e.g., *if X-then Y*) takes over and offers a completion for the not-yet-completed turn at preliminary component possible completion point using the projected final component form (Lerner, 1996a). In other words, the recipient co-opts the completion of the turn-in-progress on its speaker's behalf, which can then be accepted, rejected or modified by the original speaker (Lerner, 2004; Sidnell, 2010). Research has also shown that not only the compound turn-constructive units but also other features of interaction such as word repetitions or searches, intra-turn pauses and non-verbal behaviour can provide opportunities for turn completion by another co-participant (Lerner, 1996a). Choral co-production takes place when a participant in a conversation simultaneously co-produces the same words in a TCU more or less in the same way with the current speaker (Lerner, 2002). One social action that is achieved by the practices of anticipatory completion and choral co-production is to show agreement or affiliation with the other speaker when it is made relevant next (Lerner, 1996b, 2002, 2004; Hayashi, 2013; Sidnell, 2010; Szczepek, 2000; Van Kruiningen, 2013). Collaboratively built turns also allow recipients to demonstrate understanding of what is being said (Lerner, 2004; Sert, 2019; Sidnell, 2010; Szczepek, 2000; Van Kruiningen, 2013).

Sequence organization is a fundamental feature of talk that is examined in CA studies since turns-at-talk are produced in a sequential context and they are built in relation to what came before in the prior turn with implications for the next turns (Drew, 2013). In other words, the production of each utterance in a conversation is highly dependent on its position in a sequence. Participants initiate, develop and complete an activity and manage their interaction by means of the organization of social actions in sequences (Heritage, 1997;

Heritage & Clayman, 2010). As Schegloff (2007) defines, sequence organization refers to “the organization of courses of action enacted through turns-at-talk- coherent, orderly, meaningful successions or ‘sequences’ of actions or ‘moves’”, which are “the vehicle for getting some activity accomplished” (p.2). This is explained by Sacks (1987) with the principle of contiguity, which the participants orient to by designing their turn in connection with the immediate prior turn through, for instance, appropriate next actions and displaying this connectedness. The notion of ‘nextness’ put forward by Sacks (1992) also explains the reflexive relationship that each utterance has with the prior and the upcoming turn. The analysis of the sequencing of actions in talk facilitates an understanding of the meaning of social actions performed in social interactions (Stivers, 2013). In relation to this, Johansson (2002) states that

in interaction, every action projects the next action which in turn creates a link with the action preceding it. Every utterance derives its meaning and function according to its placement in the sequence of action. Moreover, interaction can be considered as the speakers’ interpretation of the situation and the context they are creating (p. 251).

While organizing a sequence of talk, participants in interaction orient to a shared set of interactional norms that “guide their conduct and understanding of how their conduct will be understood by others” (Stivers, 2013, p.200). Consequently, for an empirical study of social interaction and the development of a theory with respect to its organization, there is a need to reveal how social interaction is organized sequentially; that is “how turns relate to sequences; how sequences relate to activities; and how sequences and activities relate to an overall structural organization of a given interaction” (Stivers, 2013, p.208).

Although interaction always consists of an action sequence (Seedhouse, 2004), there are different forms of sequence organization such as extended telling; however, the basic unit of sequence organization is adjacency pairs, which provide a solution to the problem of what to do next in conversation (Schegloff, 2007). For instance, a question makes an answer relevant next, and an offer tends to be followed by an acceptance or a rejection. The features of adjacency pairs as the basic form of sequence organization in talk were described by Schegloff (2007, p.13) as follows:

- (a) composed of two turns
- (b) by different speakers
- (c) adjacently placed; that is, one after the other
- (d) these two turns are relatively ordered; that is, they are differentiated into “first pair parts” (FPPs) and “second pair parts” (SPPs)
- (e) the components of an adjacency pair are pair-type related

Adjacency pairs are composed of two turns, namely first pair parts (FPPs) (e.g., questions, offers, requests, invitations, advice) and second pair parts (SPPs) (e.g., answers, acceptances, rejections, declines, agreements and disagreements), which are launched by two different speakers. FPP, which initiates an exchange, makes a type-fitted SPP conditionally relevant next as a responsive action to the prior turn and recipients show their understanding of the prior turn by doing a relevant SPP (Schegloff, 2007). The absence of the relevant next action; on the other hand, becomes noticeable and is treated as accountable and sanctionable by the participants (Schegloff, 2007). Adjacency pair organization is a framework of reciprocal conduct involving both action and interpretation since participants must analyze the other's action in order to produce appropriate reciprocal action and display their understanding of one another's actions (Goodwin & Heritage, 1990).

The adjacency pairs usually have two possible second pair parts (e.g., acceptances or rejections, agreement or disagreement) and those instances are explained by the mechanism of *preference organization* in talk, which is an element of sequence organization. It explains how certain actions are designed and managed to prefer particular kinds of responses or actions in the next turn and how the recipients infer the intent of the prior speaker and analyze the initiating action to determine the relevant preference principles (Pomerantz & Heritage, 2013). Schegloff (2007) notes that the issue of preference refers to “a structural relationship of sequence parts” and is “a social/interactional feature of sequences and of orientations to them” rather than a psychological one (p.61). The concepts of ‘preferred’ and ‘dispreferred’ are, thus, social in nature (Liddicoat, 2007). Consequently, affiliative responses such as acceptance of an offer or agreement with an assessment, which build social solidarity, are preferred responses while disaffiliative responses such as rejection of an invitation, which are problematic for social relationships, are dispreferred responses (Liddicoat, 2007; Seedhouse, 2004).

Dispreferred responses are designed in different ways in order not to disrupt the relationship that exists between the speaker and the recipient and so as to minimize the degree of disaffiliation and emerging conflict (Liddicoat, 2007; Seedhouse, 2004). As a result, dispreferred responses such as “disagreements, disconfirmations and rejections are frequently performed with delays, mitigations, understated components, and accounts” as well as prefaces (e.g., agreement-prefaced disagreements), “partial or weak versions of agreements, confirmations or acceptances incorporated in them”, questioning repeats and

requests for clarification (Pomerantz, 1984a; Pomerantz & Heritage, 2013, p.214). In relation to this, dispreferred responses are likely to be more elaborated, being commonly accompanied by accounts, excuses, disclaimers and hedges as well as apologies and appreciations whereas preferred responses tend to be short and are not treated as accountable (Liddicoat, 2007; Seedhouse, 2004; Schegloff, 2007). With respect to the positioning of dispreferred responses, as Schegloff (2007) further explains, dispreferred responses break the contiguity between the first and the second pair parts through the use of such practices as (1) inter-turn gap, (2) turn-initial delay via hedges and discourse markers (e.g., “*well*”), (3) anticipatory accounts, expressions of appreciations, excuses, and (4) pro-forma agreement, which is an agreeing response that serves to delay the dispreferred one via, for instance, the “*yes, but...*” structure and where dispreferred second pair part “gets done as an exception or modification of an initial, apparently preferred response” (p.70).

As Pomerantz and Heritage (2013) explain, when a response is delayed signalling a possible upcoming dispreferred response (e.g., a potential disagreement) that is as-yet-unstated, the prior speaker can change his/her position and redesign or modify his/her initiating action or back down from it in order to elicit a preferred response. Schegloff (2007) refers to this as pre-emptive reformulation with preference reversal. For instance, in the case of yes-no questions, they are generally (re)designed and their displayed preference is reversed to enable confirming responses (Pomerantz & Heritage, 2013). What is more, preference is not only relevant to SPPs but also to FPPs as some FPP types such as criticism of co-participants (i.e., negative other assessments) and requests are dispreferred actions (Liddicoat, 2007; Schegloff, 2007). Thus, they are routinely constructed as dispreferred turns through the deployment of delaying devices as well as accounts and mitigations (Liddicoat, 2007).

As Schegloff (2007) further explains, the basic adjacency pairs can also be systematically expanded before the FPP (i.e., pre-expansion), between the two parts (i.e., insert expansion), or after the second pair part (i.e., post-expansion). Those sequence expansions usually consist of adjacency pairs in their own right, serving the base sequence in some way. Pre-expansions project some specific next action such as invitations, requests and offers and are preparatory to those projected actions to be done in the sequence, functioning as pre-invitation, pre-request and pre-offer sequences (Liddicoat, 2007; Sidnell, 2010). In doing this, they establish the necessary pre-conditions and lay the groundwork for a projected base FPP, influencing and organizing the way the subsequent talk will unfold (Liddicoat, 2007;

Stivers, 2013). As Sidnell (2010) states, pre-expansions “typically check on a condition for the successful accomplishment of the base first pair part” (p.97). Therefore, “the production of the projected base FPP is contingent upon the outcome of the pre-sequence” (Liddicoat, 2007, p.128). For example, before an invitation a speaker might check the recipient’s availability and then make the invitation based on the response s/he receives in the pre-expansion sequence.

Pre-sequences help avoid projected dispreferred sequence types and some future trouble such as conflict, disagreement and rejection as they give the speakers the chance to adjust their base first pair part or abort it (Pomerantz & Heritage, 2013). To illustrate, prior to an invitation, if a speaker learns that the recipient is busy and is not available through his/her preliminary inquiry in the pre-expansion sequence, then the invitation is unlikely to take place and the dispreferred response of a rejection is avoided (Pomerantz & Heritage, 2013). Considering the responses in the pre-sequence, they can, thus, act as a go-ahead response or a blocking response. As Schegloff (2007) describes, while go-ahead responses promote the progress of the sequence by encouraging its recipient to produce the projected base FPP, blocking responses discourage the occurrence of the base FPP when, for example, they raise the likelihood of rejection to an invitation. Therefore, these responses in the pre-expansion sequence “help the interaction avoid a sequence with a rejected base FPP” (Schegloff, 2007, p.31). Pre-sequences, thus, aim to maximize the likelihood of acceptance and minimize the likelihood of rejection by encouraging or discouraging the subsequent production of the base sequence (Schegloff, 2007). Additionally, pre-sequences can have a “pre-empting” response when the prospective recipient undertakes the projected base action (i.e., FPP) before the speaker does (Schegloff, 2007). For example, in the case of a pre-announcement, the recipient might not only claim to already know the matter at hand but also display this knowledge or make a guess by pre-emptively implementing the telling action that the pre-sequence was meant to project. Similarly, a pre-sequence that projects a request may elicit an offer from the recipient. This removes the need to perform dispreferred FPPs such as requests or telling bad news (Liddicoat, 2007).

Insert expansions initiated by prospective SPP speakers affect the progressivity of the base sequence by intervening between the first and second pair parts in the form of a sequence of its own although they are still relevant to the base sequence of action underway and mainly appear in two forms: post-first and pre-second (Schegloff, 2007). While post-first insert

expansions are other-initiated repair sequences that address some issue related to the base FPP (e.g., problems of speaking, hearing or understanding the preceding talk), pre-second insert expansions are generally type-specific and are concerned with establishing the necessary contingencies that are preliminary to a particular type of response that is to come in the base SPP (Schegloff, 2007; Stivers, 2013). In this respect, insert expansions do the interactional work that is necessary for the accomplishment of the base sequence successfully (Liddicoat, 2007).

As for post-expansion sequence, it can be explained with reference to preference organization. Schegloff (2007) states that sequences with preferred or agreeing second pair parts are “closure-relevant” and those with dispreferred or disagreeing second pair parts are “expansion-relevant”, leading to the occurrence of post-expansion sequences (p.117). There are two kinds of post-expansions: minimal and non-minimal (Schegloff, 2007). Minimal post-expansions consist of one additional single turn to a base sequence following its SPP and this turn is designed not to project any further talk within a particular sequence but to offer a minimal reaction to the base SPP by the speaker of the initiating action (Schegloff, 2007; Stivers, 2013). This means that minimal post-expansions do not initiate a new sequence, appearing in the third turn position and thus, are referred to as sequence closing thirds (Schegloff, 2007; Stivers, 2013). For instance, they include response tokens such as *oh*, *okay* and *great* or a combination of them, which usually convey that the response to the initiating action (i.e., the base FPP) was adequate and propose sequence closure (Schegloff, 2007; Stivers, 2013).

On the other hand, non-minimal post-expansion (NMP) sequences generally treat a response as insufficient for sequence closure and project further talk by initiating a new sequence due to, for example, a topical interest that invites further elaboration through the use of newsmarks (e.g., repeats, tokens such as “*really*”, *oh*-prefaced questions) or a problem of speaking, hearing and understanding occurring in the base SPP (i.e., post-second repair) (Liddicoat, 2007; Schegloff, 2007; Stivers, 2013). Therefore, NMP sequences can have various functions such as request for clarification, confirmation and elaboration as well as other-initiated repair of some aspect of a base SPP. In addition, as Schegloff (2007) points out, NMP sequences can be disagreement-implicated which is in close relation to the observation that dispreferred second pair parts are sequence-expansion-relevant. That is, NMP sequences provide a location for disagreeing with a SPP or for challenging it, which

can then be either accepted or rejected in the following turn by the base SPP speaker (Liddicoat, 2007). In the case of a rejection of disagreement with the base SPP, the post-expansion sequence is further expanded while it is briefer if the disagreement is accepted (Liddicoat, 2007). This indicates that post-expansion is “one place where the consequences of dispreferred second pair parts get played out” and where participants make an attempt to resolve the misalignment and move towards compromise as a result of an emerging disagreement and the challenging of the SPP (Schegloff, 2007, p.152). In this case, the questioned SPP can either remain or be abandoned when alternative responses are offered and the sequence is expanded through a series of SPPs. Furthermore, post-expansions can not only address SPPs but also FPPs through their “renewals, reworkings, repetitions, modifications, enhancements, inducements” and so on as a consequence of a dispreferred SPP in order to secure an aligning or a preferred response from the recipient (Schegloff, 2007, p.162).

As one of other mechanics of talk, repair is used to deal with some type of interactional trouble in speaking, hearing, or understanding the talk that threatens intersubjectivity in the talk-in-interaction (Schegloff, 1992b). As Gardner (2004) states, “generally the organization, structures, and coherence of talk are maintained, and that when that orderliness is threatened with breakdown, overwhelmingly that threat is dealt with very quickly, and orderliness is restored” (p.274). Although it delays the progressivity of the turn or sequence, repair is a locally managed, interactionally organized practice through which participants achieve and maintain intersubjective understanding within the sequential context of talk-in-interaction (Fox et al., 2013; Liddicoat, 2007; Schegloff, 1992b). In CA, a distinction has been made between initiating repair and actually performing the repair action by producing the repair solution because these two actions can be carried out by two different participants (Fox et al., 2013; Kitzinger, 2013; Schegloff, 2007). Therefore, repair can take the forms of self-repair and other-repair as well as self-initiated and other-initiated repair depending on who initiates and who completes the repair. As Schegloff (2007) explains, self-initiated self-repair occurs in the same turn when a speaker of the trouble-source or repairable cuts off his/her own talk to initiate repair and deal with a problem him/herself during the production of a turn in the form of, for example, replacement, insertion, deletion, or reordering of the units of a talk. The speaker may also initiate repair in the third turn rather than the same turn when s/he notices a trouble upon the recipient’s response. On the other hand, the repair action can be done more collaboratively whereby the recipient this time locates and indicates



some sort of a difficulty or trouble in the preceding turn and undertakes repair, resulting in other-initiated other-repair (Schegloff, 2007). However, it is more common for other-initiation of repair to lead to self-repair where the speaker of the trouble source provides the repair solution upon the recipient's indication of a problem in the prior turn, which engenders a sequence itself (Fox et al., 2013; Schegloff, 2007). In relation to this, it is found that there is a preference for self-initiations of repair over other-initiations and for self-repair over other-repair (Fox et al., 2013; Seedhouse, 2004).

As for turn design in CA, it refers to "how a speaker constructs a turn-at-talk- what is selected or what goes into 'building' a turn to do the action it is designed to do, in such a way as to be understood as doing that action" (Drew, 2013, p.132). In other words, "turns are designed to be recognizable by others as doing what they mean to be doing" (Drew, 2005, p.94). Turns constructed at talk consist of identifiable components or units (i.e., TCUs) including single words, clauses or phrases, single sentences or combination of these. As Drew (2005) explains, when a speaker designs his/her turn, s/he selects the action that the turn will perform and the details of verbal constructions (e.g., phonetic, lexical, syntactic, grammatical, and embodied) through which that action will be accomplished. Once again, this brings into view the importance of social action in conversation as it is through these actions that social life is conducted and social relationships are managed and these actions are performed through the construction of turns at talk; that is, people do things in their turns at talk (Drew, 2005).

When deciding on what action to perform, speakers try to make sense of each other's talk and thus, analyzing the turn design helps to find out the mutual understandings of the participants, their perceptions of the actions and their attempts in managing the interaction by making their verbal behaviour recognizable to others as intended. By designing their turns, speakers not only display their understanding of an immediately preceding interaction but also create a context for the upcoming talk, resulting in a collaborative achievement of actions. According to Drew (2005), CA research focuses not just on social actions but also on "participants' understandings of those actions, and particularly the basis on which such understandings were arrived at" since what connects turn design with the accomplishment of a social action is how participants come to understand the other's conduct as implementing a particular kind of action based on the practices through which turns are constructed (p.87). In this respect, turn design is closely related to the issue of meaning making. As Gardner

(2004) points out, the aim in the analysis of turn design is “to demonstrate how certain constructions are chosen to achieve particular actions, and how these choices are motivated, in part at least, by local interactional contingencies” (p.275).

The mechanics or the basic practices of talk just explained shed light on the participants’ *interactional competence* that has been discussed as a construct on its own. In the broadest sense, it refers to the participants’ ability to communicate in orderly and meaningful ways with one another (Drew, 2005) and it is an ability that develops from early childhood on (Pekarek Doehler & Petitjean, 2017). More specifically, interactional competence is a collaborative, socially constituted domain of communicative competence including interactional practices such as sequencing, turn-taking and repair, which are necessary for the accomplishment of socially relevant acts of communication (Markee, 2000). As Drew (2005) explains, interactional competence consists of the practices and devices that participants use to produce and understand conduct in talk-in-interaction and that are the focus of CA research. Those are practices that establish the highly organized structures and patterns in conversation (Drew, 2005). Pekarek Doehler and Petitjean (2017) make a similar remark with respect to interactional competence by defining it as “a set of systematic procedures through which participants coordinate their actions and deal with the local contingencies of the courses of action they are engaged in” since social actions whether it be chatting around a cup of coffee or interacting in a professional meeting require highly organized socially coordinated procedures (p.4). According to Walsh (2011), interactional competence is about more than being accurate or fluent since it involves a great deal of mental and interactional abilities such as paying close attention to the local context, listening, showing signs of understanding, attending to each other’s contributions, clarifying meaning, and repairing; in other words, managing communication and jointly constructing meaning. Therefore, interactional competence is highly context-specific and collaborative and involves both interactional and linguistic resources but with a greater emphasis on specific interactional strategies such as turn-taking, overlaps, interruptions, repair and topic management. Similarly, Wong and Waring (2010) define interactional competence as “the ability to use the various interactional resources such as doing turn-taking or dealing with problems of understanding” (p.7). Furthermore, as a phenomenon that is co-constructed by the participants, interactional competence has a local and adaptive nature, referring to the participants’ capability of continually revising and adapting their interactional practices in context-sensitive ways (Pekarek Doehler & Petitjean, 2017). Pekarek Doehler and Petitjean

(2017) describe this understanding of interactional competence as “profoundly dynamic, adaptive and situated, involving the locally contingent and recipient designed use of resources for social interaction” (p.20).

Considering the importance of interactional competence, it is through interaction that we handle our daily lives and duties and the success of our interactions depends on our level of interactional competence (Wong & Waring, 2010). As Walsh (2011) states, “effective communication rests on an ability to interact with others and to collectively reach understandings. Interactional competence, then, is what is needed in order to ‘survive’ most communicative encounters” (p.159). Although technical skills and specialized knowledge are needed for some settings, interactional competence “represents a general instrument for problem-solving, participation, recognition and success across a large variety of settings” and encompasses a set of practical procedures (e.g., turn-taking, requesting, telling stories, etc.) through which “participants in an interaction coordinate their actions, guide each other, accomplish roles and relationships and give their encounter features of typicality, objectivity and normality” (Pekarek Doehler & Petitjean, 2017, p.7-8).

In the context of institutional settings, as a central component of the wider social abilities, interactional competence enables people to pursue educational and professional goals and gain access to various institutional worlds, playing an important role in issues such as workplace socialization, identity construction, learning, professional integration and involvement in a community of practice (Pekarek Doehler & Petitjean, 2017). Due to the interactional demands and challenges presented by institutions (e.g., multitasking, adjusting to multiple types of participants, their roles, duties and entitlements, changing participation frameworks, time pressure), interactional competencies constitute an important resource to deal with those demands since they are “central instruments for participation, task accomplishment, self-presentation as well as membership” (Pekarek Doehler & Petitjean, 2017, p.10). This means that young people need to adapt their existing interactional competencies and elaborate new ones to deal with a variety of new interactional situations when they move into the world of work (Pekarek Doehler & Petitjean, 2017).

Regarding the role of CA in studying interactional competence, Wong and Waring (2010) state that “conversation analysis offers a wealth of knowledge that can make our understanding of interactional competence more specific, more systematic, and more

pedagogically sound” (p.8). CA allows access to interactional practices, which we use to engage in social interaction. Since interaction is made possible through certain shared communicative competencies (e.g. knowledge of structures, patterns, norms etc.), by studying conversation, such fundamental competencies that establish mutual understanding and make conversation between members of a culture possible are brought into view (Drew, 2005). As Heritage and Atkinson (1984) point out, CA aims to describe and explicate the competencies that speakers use and rely on in participating in a socially organized interaction by unveiling “the procedures by which conversationalists produce their own behaviour and understand and deal with the behaviour of others” (p.1). Pekarek Doehler and Petitjean (2017) further acknowledge that CA is “one of the most prominent lines of research that informs us about the workings of interactional competence” (p.7) and that current research needs to reveal what interactional competencies are solicited and how they are deployed in various contexts and the challenges that emerge for people interacting in those contexts. According to them, CA provides a strong basis for understanding the systematic procedures or methods (e.g., turn-taking, repairing, opening or closing a conversation) that the participants use for organizing and engaging in social interaction as part of their interactional competence, treating them as full members and competent participants. Additionally, as Seedhouse (2011) indicates, in line with the definitions of interactional competence, CA treats interactional competence as variable and co-constructed by participants in interaction rather than as fixed and static and offers a means of examining how such competence is co-constructed by the participants in particular contexts. Since CA is concerned with publicly observable phenomena, it views competence as involving the use of situated practices rather than deriving from psycholinguistic models of learning processes and knowledge structures (Schegloff et al., 2002).

With respect to the criticisms made of CA approach, CA was criticized for the inability to generalize findings to other research contexts as in many qualitative studies because CA focuses on a very specific context (Walsh, 2011) and for having insufficient instances of data with a focus on micro incidents and single cases and not enough reported frequencies (Psathas, 1995). In response to these criticisms, Walsh (2011) puts forward that as in the case of all qualitative research, the aim of CA research is not to generalize but “to promote understanding and facilitate replication to another context” (p.88). The aim is thus, to report trends, patterns and tendencies rather than absolute conclusions. As Gardner (2004) also indicates, CA is not “a search for generalizations, but a description of how an interaction

works in and of itself” and “of what is going on uniquely” (p.266). That is, by examining instances of interaction, CA aims to explain the methods that members of a culture use in accomplishing actions (i.e., methods for interpreting their own and other participants’ behaviours) and to present the sequential and interactional features of those methods rather than categorizing their actions and finding the right names with which to label them (Pomerantz, 1990). In addition, CA provides accounts which are both particularized and generalized because although specific features of individual contexts are described, the sequential pattern identified based on those specific features may present a general picture, contributing to the external validity of CA studies (Hutchby & Wooffitt; 1998; Seedhouse, 2005). Therefore, as Seedhouse (2005) points out, CA studies “work on the particular and the general simultaneously; by analysing individual instances, the machinery that produced these individual instances is revealed” and this makes it possible to provide a generalizable description of the interactional organization of social actions which are seen as the products of a machinery (p.256).

What contributes to the generalizability of findings is also (1) the accumulation of empirical findings that have been revealed over the years with respect to interactional practices organizing social action (Gardner, 2004; Heritage & Clayman, 2010), (2) the comparative approach used where individual instances are compared to each other to identify a recurrent phenomenon, which is then compared in different settings to find out its setting-specific features, parallel instances as well as deviant cases, to make modifications and additions to the underlying pattern where necessary until a coherent pattern is obtained and to generate general formulations of ‘devices’, ‘organizations’, or ‘systematics’ (Arminen, 2005; Drew, 2003; Ten Have, 2007), and (3) the collection-based approach that has been adopted in CA whereby “the conversation analyst must collect multiple instances of a phenomenon in order to discern the generic, context-independent properties of a practice” through the analysis of the particularities of the single cases in the collection (Sidnell, 2013, p.78). However, as Sidnell (2013) notes, while using collections, it is important for the analyst to remain accountable to the particularities of each individual case. In other words, no rigorous claims are made about the extent to which particular interactional methods and features are used and the users or the circumstances where they are used (Pomerantz, 1990). According to Pomerantz (1990), CA studies lead to the identification of “features of interactants’ interpretive work that otherwise are undefined, hazy, and undifferentiated” and “offer them as proposals to the community of social scientists” so that once they are differentiated and

clarified, their generality can be determined by, for instance, obtaining different collections of data from different or similar environments (p.234). Moreover, in response to the criticism of CA for insufficient reported frequencies; in other words, in relation to the issue of quantification, Seedhouse (2005) points out that as a qualitative methodology that aims to develop an emic perspective, “quantification is generally of peripheral interest to CA practitioners” in the sense that although it is not prohibited in CA when, for example, isolating an interesting phenomenon, it should not divert the attentions of CA researchers from detailed analysis of individual instances (p.259).

The validity of CA studies is explained in terms of the next-turn proof procedure, which refers to the recipients’ responses in the next turn as a source of evidence for the analysis of the immediately prior turn and which is used as the basic tool by the analysts to develop an emic perspective (Sacks et al., 1974; Seedhouse, 2004; Sidnell, 2013). This means that the participants’ display of their own understanding of prior turns and their own orientations to the structures while constructing their turns becomes available to the analysts as well in explaining how a particular practice was understood and interpreted by the other co-participants and in revealing the patterns of talk (Sidnell, 2013). Therefore, as Sacks et al. (1974) explain, “the display of those understandings in the talk of subsequent turns affords both a resource for the analysis of prior turns and a proof procedure for professional analyses of prior turns” (p.729). In CA, the next-turn proof procedure provides one basic kind of evidence based on which analysis can be carried out and CA findings can be tested but there are also other kinds of already available data-internal evidence from the surrounding talk or within the same turn that the analysts rely on in developing their analysis (Sidnell, 2013). In relation to this, CA findings have a demonstrable validity since CA deals with what is made relevant in and through interaction through the participants’ situated orientations and meaning-making, disregarding any pre-determined categories or codes, existing theories of language, society or psychology as well as contextual features and cultural, social or personal aspects that are external to the interaction (e.g., participants’ social status, gender) and that are not displayed in the details of the interaction (Arminen, 2005; Seedhouse, 2005). Seedhouse (2005) considers such an emic perspective provided by CA as contributing to the internal validity of CA studies because “participants document their social actions to each other in the details of the interaction by normative reference to the interactional organizations” and the analysts make no claim beyond what is demonstrated by these interactional details, which provide a justification for their claims (p.255).

The video and audio recordings themselves contribute to the validity and reliability of CA studies since they provide all the details of talk-in-interaction as “naturally occurring empirical materials” and can be examined repeatedly (Heritage & Atkinson, 1984, p.4; Heritage & Clayman, 2010). The naturally occurring talk recorded in its authentic social setting makes CA studies ecologically valid as it enables the findings of CA studies to be directly applicable to people’s everyday life and gives CA researchers the opportunity to refer to the same interactional organizations that the participants are using in describing how they perform their social actions in and through talk (Seedhouse, 2005). Besides opportunities for repeated and detailed examination of everyday interactional events, the availability of recorded data along with the analyses allows the public to have direct access to the data, evaluate the findings of CA studies and check the validity of the claims made about the data, making the analysis process transparent for readers (Heritage & Atkinson, 1984; Heritage & Clayman, 2010; Seedhouse, 2005). As a result, the recorded data becomes available for “public scrutiny in a way that further minimizes the influence of individual preconceptions” (Heritage & Atkinson, 1984, p.4) and paves the way for the reproducibility and replication of findings (Sacks, 1992; Seedhouse, 2005). Therefore, it is a standard practice for CA researchers to organize data sessions where they view and listen to segments of recorded data repeatedly, discuss them in detail and share their observations or analyses of the same data fragment (Seedhouse, 2005). Furthermore, Seedhouse (2005) explains the reliability of CA studies in terms of the selection of what is recorded, the technical quality of recordings and the adequacy of transcripts.

### **2.6.1. Institutional Talk**

In CA, conversation is often referred to as “talk-in-interaction” as the umbrella term since there are different kinds of talk or speech exchange systems across a range of settings from casual or ordinary (e.g., conversation between friends) to institutional contexts (e.g., classrooms, doctor-patient consultations, courtroom interactions) (Wong & Waring, 2010). Ordinary conversation is “the kind of casual everyday talk that typically occurs between friends and acquaintances” and is “considered to be the default speech exchange system in all talk-in-interaction” (Markee, 2007, p.1017). In Drew and Heritage’s (1992) terms, it is the predominant and primary form of interaction through which socialization proceeds. This means that the sequential, turn-taking, and repair practices of ordinary conversation undergo certain structural modifications in other speech exchange systems (Markee, 2007;

Zimmerman & Boden, 1991). The basic set of interactional norms and resources provided by the primordial ordinary conversation offers a reference model or a comparative measure for the description of other forms of social interaction, which are regarded as transformations of ordinary conversation (Drew, 2005; Mondada, 2013). Thus, institutional talk as a different form of talk is also claimed to derive from ordinary conversation in terms of its organization and the types of actions employed (Gardner, 2004). This distinctiveness of institutional interaction can, thus, best be explained by considering the differences that have been discovered with respect to ordinary conversation and institutional talk (Drew & Heritage, 1992).

In ordinary conversation, turn taking, turn size, content and type are not predetermined, are free to vary and locally managed (Markee, 2005; 2007; Zimmerman & Boden, 1991). That is, it includes a wide range of rules and practices as well as indefinite inferential frameworks (Heritage & Clayman, 2010). It is also preferred that turn length is kept to a minimum and repair is self-initiated and self-completed (Markee, 2007). On the other hand, talk in institutional setting is viewed as goal-oriented and task-related, and the organization of the interaction is related to the institutional goal, which constrains the actions and participants' interactional contributions (Drew & Heritage, 1992; Gardner, 2004; Heritage, 2005; Heritage & Clayman, 2010; Seedhouse, 2011; Walsh, 2011) and makes the preference structure and the distribution of members' sequential, turn-taking, and repair practices quite different from ordinary conversation (Markee, 2007). In other words, in institutional contexts, participants shape their conduct by reference to institution-specific goals and constraints, which contribute to the distinctive, specialized, restricted and recurrent character of institutional interaction (Drew & Heritage, 1992; Heritage & Clayman, 2010; Zimmerman & Boden, 1991). Thus, the goal-oriented activity creates the context and makes institutional talk more ritualized and formalized than ordinary talk although participants use the same basic mechanisms (Gardner, 2004; Markee, 2005). In this respect, institutional talk is considered to be "a simplification of ordinary conversation in terms of the organization of the speech exchange system" and thus, is culturally and socially restricted (Gardner, 2004, p.264). As Heritage (1984a) indicates, institutional interactions involve reduced and specialized versions of the practices (e.g., turn-taking system, sequence types) described in ordinary talk.



With respect to another difference between ordinary conversation and institutional talk, certain actions that are avoided in ordinary conversation might be promoted in institutional contexts or conversely, those that are encouraged in ordinary conversation might be interpreted as threatening in specific institutional interactions (Drew & Heritage, 1992). For instance, in legal, medical and news-interview settings, one institutional constraint put on professionals is the withholding of expressions of surprise, sympathy, or agreement in response to lay participants in order to display epistemic caution and professional neutralism; however, this would normally be interpreted as disaffiliative in a conversational context (Drew & Heritage, 1992). This further explains the distinctive features of interactional inferences in institutional orientations to talk at work and shows how institutional talk is to be understood in terms of institution and activity specific inferential frameworks (Heritage & Clayman, 2010). Drew and Heritage (1992) summarize the features of institutional talk as in the following:

- (1) Institutional interaction involves an orientation by at least one of the participants to some core goal, task or identity (or set of them) conventionally associated with the institution in question. In short, institutional talk is normally informed by *goal orientations* of a relatively restricted conventional form.
- (2) Institutional interaction may often involve *special and particular constraints* on what one or both of the participants will treat as allowable contributions to the business at hand.
- (3) Institutional talk may be associated with *inferential frameworks* and procedures that are particular to specific institutional contexts (p.22).

Regarding the institutional context, rather than the institution providing a context for a particular type of talk, CA views it as being built by the participants in and through their talk by means of which they perform tasks and activities in pursuit of institutional goals and manifest their institutional conduct (Drew & Heritage, 1992; Heritage, 1997; 2005). CA does not start with an a priori categorization of context which would impose a set of pre-determined constraints on institutional interactions but instead, treats context as being continuously created, maintained, altered, and reflexively achieved locally by the participants in and through the actions they perform as they sequentially organize talk and orient to its institutional features (Heritage, 2005; Heritage & Clayman, 2010; Mondada, 2013). Therefore, CA addresses context in terms of, for instance, institutional resources, constraints, and specialized knowledge as it is demonstrably made relevant in interaction for the implementation of actual institutional practices.

Drew and Heritage (1992) further make a distinction between formal (e.g., classrooms, courtrooms, news interviews) and relatively less formal (e.g., medical, social-service, business, unmoderated team meetings) forms of institutional interaction. Institutional talk in formal settings is characterized by, for instance, specialized and strongly constrained turn-taking systems, actions and topics that are organized in explicit and predictable ways (e.g., action pre-allocation or mediated turn-allocation) as in courtroom, chaired meeting and teacher-fronted classroom interactions (Drew & Heritage 1992; Heritage & Clayman, 2010). In those settings, participants implement a set of interactional practices involving reductions and specializations of available conversational options as well as restrictions in their contexts, and organize their conduct recurrently and pervasively over the course of interactional events (Drew & Heritage, 1992; Heritage, 2005). This differentiates a specific form of institutional talk from other institutional interactions as well as mundane conversation, contributing to “a unique ‘fingerprint’ for each institutional form of interaction” (Drew & Heritage, 1992, p.26). In addition, interactions in formal institutional settings

are culturally variable; they are sometimes subject to legal constraints; they are always vulnerable to processes of social change; they are discursively justifiable and are often justified by reference to considerations of task, efficiency, fairness, and so on in ways that the practices making up the conversational ‘bedrock’ manifestly are not (Drew & Heritage, 1992, p.26).

On the other hand, in non-formal or less formal institutional settings, the interactions generally take place in private rather than public contexts and there is considerably less uniformity in the patterns of talk-in-interaction since “there is room within them for considerable negotiation and/or stylistic variation as to how they will come to be managed” (Drew & Heritage, 1992, p.28). Although the talk in these less formal settings is still institutional with a focus on official task-based or role-based activities, there are no rigid boundaries between these forms of institutional talk and ordinary conversation especially in terms of turn-taking procedures (Drew & Heritage, 1992). In other words, the interactional practices are more complex and variable in terms of form and frequency even though the organization of sequences and turn design within those sequences are found to have systematic aspects contributing to the institutionality of how such interactions are managed due to the goal orientation, special constraints and inferences imposed by the institutionalized set of norms (Drew & Heritage, 1992; Heritage, 2005). In this study, too, the examined interactions between testers and teachers in reviewing language tests in a

higher education setting are of the institutional nature taking place in a private context that is relatively less formal and thus, the interactants have certain roles and goals. This study, thus, aims to reveal the ‘institutionality’ of those language test IR interactions by revealing how the teachers orient to the institutional goals, tasks (i.e., test materials), constraints, and roles and responsibilities in and through the moment-to-moment talk as the item reviewing activity is “talked into being” (Heritage, 1984a, p.290).

In CA, one of the important developments has been the emergence of a corpus of studies concerned with talk in institutional settings such as medicine, education, law, business and news media where the daily activities of professionals and organizations are conducted and managed through talk-in-interaction (Drew & Heritage, 1992; Heritage, 1997; Heritage & Clayman, 2010). This has led to the concepts of ‘institutional talk’ or ‘institutional interaction’ and ‘institutional CA’, which builds on the basic findings about talk “as a means to analyze the operations of other social institutions in talk” (Drew & Heritage, 1992; Heritage & Clayman, 2010, p.16). In the context of institutional talk, CA studies focus on the participants’ orientations to not only the interactional practices but also the institutional character of their talk (e.g., specific institutional identities, tasks and constraints); in other words, “the institution as it is managed, reproduced, and eventually transformed in and through interaction” (Drew & Heritage, 1992; Heritage, 1984a, 2005; Heritage & Clayman, 2010; Mondada, 2013, p.2). As Heritage (1984a) states, “it is within these local sequences of talk, and only there, that these institutions are ultimately and accountably talked into being” (p.290). Those studies have provided the distinctive forms of participation and highly specialized forms of activity that arise in these specialized organizational environments and built a foundation for understanding certain aspects of ‘institutionality’ and the situated accomplishment of institutional tasks and activities (Heath & Luff, 2013; Heritage, 1997).

The corpus of research concerned with institutional talk also includes ‘workplace studies’ which focus on the interplay of work, interaction and the use of technology in a range of organizational environments (Heath & Luff, 2013). They address how the use of tools and technologies ranging from simple artifacts (e.g., written documents, whiteboards) to complex computer systems is embedded within the collaborative production and interactional accomplishment of specialized tasks and activities in institutional environments and how they structure the ongoing activity (Heath & Luff, 2013; Mondada, 2013). CA has, thus, allowed access to the interior processes and practices of institutions (Drew & Heritage,

1992) through which they “recognize and make recognizable the skilled, expert, specialized conducts and reasonings in which they are engaged” (Mondada, 2013, p.4).

Schegloff (1992a) mentions two specific features of talk in institutional occasions which the analysts should demonstrate: (1) relevance, and (2) procedural consequentiality. With relevance, he refers to how the participants orient to the relevances of a work setting or a specific task through the details of their talk for the interactional accomplishment of a conduct. In relation to this, Schegloff (1992a) states that the analysis of institutional interaction aims to “show how the parties are embodying for one another the relevancies of the interaction and are thereby producing the social structure” (p.110). As for procedural consequentiality, Schegloff (1992a) refers to how the participants’ demonstrable orientation to the institutional context being formulated is procedurally consequential to the talk, that is, “for the shape, form, trajectory, content, or character of the interaction that the parties conduct” (i.e., what actually happens in the talk) (p.111). Therefore, as Drew and Heritage (1992) point out, studies of institutional interaction describe “the institutional nature of the tasks and relevances that inform conduct in a variety of work settings” and examine “how these relevances are established and how specific tasks are discharged through such conduct. And they depict the consequences which both may have for the character of the interaction and its outcomes” (p.4). Put differently, according to Heritage and Clayman (2010), institutional interactions have causes and consequences in the sense that external factors might influence the deployment of certain interactional practices or give rise to new ones and the use of particular interactional practices would in turn affect the outcomes of interaction (e.g., decision-making, attitudes) and the effectiveness of activities. As Schegloff et al. (2002) indicate, it is not necessarily the physical environment where talk is conducted but rather the relevance and procedural consequentiality of the institutional character and context to the participants as manifested in talk that determine the institutionality of talk and thus, “treating episodes of talk in interaction as ‘institutional’ involves showing how that institutional character is embodied- is ‘done’- in the details of talk and other conduct” (p.11).

Six overlapping dimensions of interaction have been proposed to examine the institutional character of a talk-in-interaction and to show how participants orient to the institutional context of their talk as they organize their conduct: (1) turn-taking organization, (2) overall structural organization, (3) sequence organization, (4) turn design, (5) lexical choice, and (6) forms of interactional asymmetry (Drew & Heritage, 1992; Heritage, 1997; 2005; Heritage

& Clayman, 2010). These dimensions of interaction constitute the layers of organization within institutional talk as one builds upon the other; that is, lexical choice contributes to turn design and the implementation of actions within turns and turn design contributes to sequence organization, which then creates the phases of interaction that make up the overall structural organization and all these activities operate through the turn-taking system (Heritage & Clayman, 2010). As Heritage and Clayman (2010) further state, analyzing the layers of organization in institutional talk leads to an understanding of the constraints under which professionals perform their tasks, how they continuously build their institutional contexts within a framework of agency and constraint, and how they situate themselves within the institutional setting in question. Besides the basic mechanisms of talk that have already been described for the analysis of ordinary talk in CA (i.e., turn-taking, sequence organization, repair, turn design) (See Section 2.6), the specific dimensions of overall structural organization, lexical choice and forms of asymmetry have also been considered for the analysis of the institutionality of a talk-in-interaction occurring in institutional settings.

Many kinds of institutional interactions are “characteristically organized into a standard ‘shape’ or order of phases” that are often task-related and that take place recurrently across ranges of instances, persons and so on with a joint orientation towards them (Drew & Heritage, 1992, p.43). This is because professionals tend to develop standard patterns of sequences for managing their routine tasks in and through interaction (Drew & Heritage, 1992). Thus, as CA research has shown, the recurrent stages of activities are an achieved order established by the participants through practical reasoning and coordinated actions (Kasper & Wagner, 2014). As Robinson (2013) further indicates, in institutional settings, participants generally orient to the activities they perform in and through interaction as having an overall structural organization and conversation analysts aim to demonstrate how such an underlying interactional pattern is both relevant to and procedurally consequential for the production and understanding of institutional activities. Activities often consist of more than one base adjacency pair sequence of action and the overall structural organization of those activities provides an interactional coherence that “shapes and constrains participants’ production and understanding of behaviour in interaction” (Robinson, 2013, p.278). In addition, each stage in the overall structural organization of an interaction pursues a specific goal or a sub-goal and is jointly constructed by the participants in the process of completing the task at hand in and through interaction (Heritage, 2005; Heritage & Clayman, 2010). Thus, overall structural organization enables both progressivity of an activity through

the completion of its components and the projectability of the actions within an interactional organization, which are organized relative to each other and which have a reflexive relationship (Robinson, 2013).

In some contexts, the oriented phases may be prescribed, for instance, through a written schedule while in others they may be locally managed and thus, may be subject to variations with certain stages being optional in some instances depending on a range of contingencies (Drew & Heritage, 1992). In relation to this, as Heritage (2005) points out, the discovery of an order of phases does not mean that each of those phases will always take place in every institutional interaction of a particular type because “the parties may well reopen sections and reinstate task orientations that they had previously treated as complete” (p.122). On the other hand, the kind of standardized and repetitive overall structural organization is likely to be found in highly focused mono-topical task-oriented interactions such as bureaucratic and service encounters (Heritage, 2005). As a result, rather than finding invariance in the order of phases, the purpose of describing overall structural organization is to “identify task orientations which the participants routinely co-construct in routine ways” and hence, it is not a “framework -fixed once and for all- to fit data into” (Heritage, 1997, p.229).

According to Heritage (1997) and Heritage and Clayman (2010), doing an analysis of the overall structural organization of an institutional interaction enables the analysts to find out (1) the number and kinds of topics discussed or dealt with, (2) the significant stages in the participants’ construction of the tasks and goals of the interaction as they move towards the completion of the work, (3) how the participants develop a joint sense of the task at hand and the roles they play within each section or phase, and (4) the boundaries that the participants agree about as they move from one each section to another (i.e., how they make a transition from one section to the next). Additionally, such a description of overall structural organization of an activity allows participants to accomplish successive tasks in an efficient and timely manner, which is especially important for the design of training manuals for novices in institutional settings (Kasper & Wagner, 2014).

As for the dimension of lexical choice, as Drew and Heritage (1992) explain, the vocabulary used in institutional talk makes reference to specialized knowledge and institutional identities as well as roles and circumstances and is selected in context-sensitive ways. Hence, institution-specific registers are associated with epistemic access and authority

(Kasper & Wagner, 2014). The choice of particular words or phrases can convey a participant's stance towards an interactional context (Heritage, 2005). To illustrate, as Sacks (1992) first pointed out, participants may use the inclusive 'we' rather than 'I' to refer to themselves as members of an organization. Through the use of an institutional self-referring 'we', they "invoke an institutional over a personal identity, thereby indicating that they are speaking as representatives, or on behalf, of an organization" (Drew & Heritage, 1992, p.30). Similarly, participants can choose to use institutional euphemisms in order to reduce the potential for conflict or disagreement in, for example, bureaucratic interactions (Heritage, 2005).

With regard to the dimension of social epistemology and forms of interactional asymmetry in analyzing institutional talk, it is explained in relation to professional cautiousness in interaction. Drew and Heritage (1992) make the remark that professionals in institutional interactions most often "design their talk so as to maintain a cautiousness, or even a position of neutrality with respect to their co-participants" (p.47). They, thus, tend to avoid taking firm positions and make their claims in cautious ways (Heritage, 1997). However, they at the same time convey a specific expert authority by making claims within a relevant knowledge domain that is based on distinctive, specialized and superior knowledge (Heritage, 1997). Furthermore, institutional interactions are characterized by the asymmetrical relationships between participants especially if they take place between professionals and lay people; however, what defines this asymmetry is the participants' roles and status, their knowhow about the interaction and the institution where it takes place as well as the "differential distribution of knowledge, rights to knowledge, access to conversational resources, and to participation in the interaction" (Heritage, 1997; Drew & Heritage, 1992, p.49). In relation to this, institutionalized expert knowledge on its own may not be enough in institutional interactions as professionals should also be entitled to that knowledge and have rights of access to it (Heritage, 1997). As Heritage (1997) indicates, in institutional interaction, "there is a direct relationship between interactional roles and tasks on the one hand and discursive rights and obligations on the other", which causes the institutional representatives to direct the interaction in ways that are different from ordinary conversation (e.g., when a topic is concluded, what the next topic will be, how the topic is to be shaped) (p.237). CA research focuses on the ways participants demonstrably orient to role-related asymmetries and the consequences this has for talk-in-interaction rather than relying on exogenous variables for the interpretation of asymmetries in conduct (Drew & Heritage, 1992).

With the application of CA to talk in various institutional settings and to fields of study with different theoretical bases as well as the practical consequences and implications it has brought in, researchers have started to discuss the notion of applied CA (Antaki, 2011; Kasper & Wagner, 2014; Mondada, 2013; Ten Have, 2007). As Ten Have (2007) states, CA-inspired studies have been carried out “to make social life ‘better’ in some way, to provide data-based analytic suggestions for, or critiques of, the ways in which social life can be organized” (p.174). This is because CA can shed light on the workings of a society’s institutions and how they manage to carry out their routine tasks successfully through talk, and further go beyond this to suggest improvements in their activities or services and to solve their problems in the case of, for example, a failed medical procedure or an unsuccessful teaching practice (Antaki, 2011). Therefore, in applied CA, studies focusing on specific organizational features and the systematic analysis of institutional practices are not only undertaken for informative purposes and consciousness-raising but they might also aim at social intervention for the training, counselling and empowerment of professional people (e.g., L2 teacher education) as well as for change in professional practices (Mondada, 2013; Ten Have, 2007).

Even if those studies of institutional talk are not specifically concerned with the training of practitioners, they can still contribute to their professional development besides enriching the scientific literature (Mondada, 2013). With such an implication, the studies of institutional interactions can enable the professionals to reflect on and evaluate their work practices by watching and discussing segments of video recordings accompanied by their transcripts in organized sessions (Arminen, 2005). This is because talk-at-work is a site for not only the use but also the development and assessment of interactional competencies demanded of professionals by the institutional settings (Pekarek Doehler & Petitjean, 2017). With regard to CA’s potential for intervention in institutional contexts, Heritage and Clayman (2010) state that “showing participants recorded data and pointing out the relevance of particular interactional practices...could be revelatory for participants and introduced new potentials for institutional reflexivity and organizational change” (p. 281). To illustrate, Sert (2015) proposed a CA-informed framework named IMDAT, which is a reflective teacher education model that can be implemented in both pre-service and in-service teacher education programs. Such a model integrates conversation analytic findings on L2 classroom interaction into teacher education whereby those findings are used as resources and audio-visual materials for training teachers in their pedagogical practices and



requires teachers to not only record their teaching practices but also critically reflect on those recorded classrooms in collaboration with their trainers using the principles of CA (Sert, 2015).

Other than an institutional and an interventionist focus, what has led to the emergence of applied CA is also (1) the insight CA has brought to different fields of study such as pragmatics, linguistics, psychology and anthropology by respecifying their basic intellectual foundations and understandings (i.e., foundational applied CA), and the use of CA (2) for investigating and understanding macro-societal problems and issues such as inequality, power, conflict, and gender (i.e., social-problem applied CA), (3) for understanding and treating communication problems and disorders (i.e., communicational applied CA), and (4) for medical diagnosis by correlating speech features with some underlying organic or psychological disorder (i.e., diagnostic applied CA) (Antaki, 2011). Finally, considering the role of CA in applied linguistics, CA-informed research has been carried out in applied linguistic research contexts such as intercultural communication, talk in educational institutions, language pedagogy, interactional competence of second language learners as well as implications for the design of language teaching tasks, materials, and assessment tasks (Kasper & Wagner, 2014; Schegloff et al., 2002; Sert & Seedhouse, 2011).

In this study too, CA is used in an applied sense by examining talk in an educational institution, offering a systematic way of scrutinizing the interactional features of language test co-construction and raising teachers' awareness with respect to the practice of syllabus-based language test item reviewing, which is a routine task commonly performed at English preparatory schools in higher education settings in Turkey. The use of CA in this study also has important implications for the field of language assessment and testing by respecifying and developing a new understanding of the practice of language test IR that has been discussed in testing literature as one of the stages of language test development and for the training of teachers in their language test construction practices. The careful analysis of everyday activities will enable the practitioners to 'deconstruct' the taken for granted within their professional settings and create a room for training and improvement. The findings of this micro-analytical study can, thus, shed light on the needs and orientations of teachers and trainers and be integrated into both pre-service and in-service L2 teacher education programs, helping teachers develop the necessary professional competencies to cope with the demands of language test IR practice and to understand the relation between test

construction and interaction in an institutional context. According to Van Kruiningen (2013), those studies that really focus on the micro level of professional interaction can provide a detailed view on the ways in which such a quality can be realized; on how shared knowledge is realized interactionally and procedurally; on how teachers actively use and reshape language structure in professional conversations in a goal-oriented and collaboration-oriented fashion (p.119).

### **2.6.2. Conversation Analytic Studies on Teacher-to-teacher Collaborative Workplace Interactions**

To the best knowledge of the researcher, no conversation analytic studies have been carried out on teacher-to-teacher interactions in relation to language test construction within the institutional context of an educational setting. However, of the CA studies on teacher-to-teacher interactions, a recent study by Sandlund and Sundqvist (2019) was conducted in the context of language assessment. Focusing on rater interactions and using a CA approach, they examined how groups of L2 English teachers collaboratively assessed learners' interactional competence in a paired L2 speaking test and how they displayed their understanding of interactional competence. In doing this, the study specifically focused on how the teachers (i.e., raters) oriented to particular learner contributions in the paired test interaction (i.e., reported speech turns) in presenting their arguments for scoring. The raters' reported turns showed what was salient to them with regard to learner performance in the test talk and served to support their arguments, contributing to the group decision-making. This study has, thus, demonstrated the collaborative assessment practice as a social, interactional process and revealed the difficulty in scoring interactional competence based on a fixed assessment criteria. Consequently, rater interactions were regarded as significant sources of information for the development of speaking tests and scoring criteria.

On the other hand, a few studies have been undertaken to examine the micro-analytic details of other kinds of collaborative workplace interactions taking place between teachers in schools. They include institutional goals and tasks such as redevelopment of a university course (Van Kruiningen, 2013), curriculum development (Barnes, 2007), discussion of unsatisfactory school marks (Verkuyten, 2000), review of rules of a system set up for managing student behaviour (Baker, 1997), and construction of a school-university partnership (Hall & Danby, 2003). Although they have a different purpose in terms of task accomplishment, post-analytic connections can be made in the CA findings of those studies

and the current study, which would shed light on teachers' collaborative task-oriented interactions with their colleagues in educational settings. The use of CA in those studies enabled the researchers to uncover and focus on different aspects of interactional patterns in workplace-based teacher meetings and encounters. They have, for instance, presented an analysis of interactive turn completions (Van Kruiningen, 2013), formulations (Barnes, 2007; Van Kruiningen, 2013), accounts and explanations (Baker, 1997; Verkuyten, 2000) as well as the use and construction of membership categorization devices such as "being an expert" (Baker, 1997; Hall & Danby, 2003).

Using CA, Van Kruiningen (2013) examined how four university lecturers established and managed intersubjective understanding and solved teaching problems interactionally in an inter-professional meeting about the redevelopment of a Biology course. In this educational meeting, the lecturers aimed to increase students' involvement in the course in question and to provide more space for their academic communication skills. Van Kruiningen (2013) specifically focused on the ways the lecturers jointly constructed a line of reasoning in *interactive turn completions* as the preliminary component of the first speaker's utterance triggered turn completion by the second speaker (i.e., compound turn constructional units). With such interactive turn completions, which were followed by acceptance and confirmation, they showed agreement, collaborative alignment and a shared knowledge base and understanding but also went beyond this to generate insights in the emerging turn. For instance, interactive turn completions emerged in the participants' elaborate collaborative search for an ideal grading system for the course and fulfilled the function of not only acknowledgement but also communicating an insight and collectively constructing a line of thought or an argument by guiding it in another direction. Eventually, uncertainties were clarified and mutual consensus was established in relation to the usefulness of grading and the ultimate assessment model. The participants also developed a sense of ownership with regard to the propositional content of the collaboratively constructed turn as their completions were accepted by the first speaker, making it easier to solve the problem and to achieve common ground in the group.

In the same study, Van Kruiningen (2013) also analyzed the use of *formulations* by successive speakers to elaborate on their interpretations of prior turns in a sequence expansion. With these formulations, they were found to "select and reformulate the gist of the foregoing and mark this element as a relevant topic for further conversation, as a way to

redefine, supplement or adjust the content of the foregoing turn” (Van Kruiningen, 2013, p.111). In this way, they provided a more concrete elaboration of the previous contributions and made them more accessible and conceivable as they collaboratively explored the possibilities of the new course plan. To illustrate, in the episode presented in the study, the idea about functions of assessment forms to be used in the course gradually took shape in a range of subsequent supplementary formulations. In doing this, the participants delivered a specification, transformation and elaboration of the initially proposed idea and led to collaborative reasoning in a goal-oriented manner. As a result, the use of an assessment form was regarded as useful due to its functions in informing the students about how they did, the criteria they met and why they performed the way they did. The lecturers collaboratively formulated those functions and contributed to the development of the concept-under-construction.

With regard to the use of formulations, Van Kruiningen’s (2013) study also revealed that formulations displayed the participants’ shared understanding and agreement with the prior speaker, functioning as affiliative moves. They appeared in topical sequence expansions and were not closure relevant; however, this was not because they were treated as dispreferred responses but because of the goal-oriented character of the interaction and the need to develop new views and possible solutions on the problems discussed. Thus, formulations in sequence expansions in this study exhibited an orientation toward alignment and consensus and thus, the construction of shared understanding. Finally, in contrast to other institutional settings where formulations were found to be the product of asymmetrical relations, in this study, they constituted an equal input of all participants with an orientation toward a reflective and collaborative process. Overall, Van Kruiningen’s (2013) study offers insight into “the ways participants work toward solutions for educational problems step by step, turn by turn, with input from all participants, where mutual expertise is made relevant in an elusive manner” (p.119). Consequently, adopting a situated perspective on educational change processes, she concludes that the fine-grained analyses of teachers’ interactional practices with respect to their practical concerns can raise awareness of the ways problems and their solutions are locally organized in and through interaction and how content, tasks and roles are negotiated in subtle ways in those interactions.

Barnes (2007) also examined *formulations* in her study on the facilitation of common agreement and shared understanding in teachers’ curriculum development meeting

interactions from a medical school. The teachers' tasks included, for instance, providing suggestions for supplementary case units to be included in the syllabus. Barnes (2007) defines formulations as "types of repeat utterances that are designed to be recognized as linked to previous discussions but completely gloss the preceding talk" (p.273). Using CA, she focused on certain subclasses of formulations such as gist formulations (i.e., summarizing the gist of a stretch of talk) and upshot formulations (i.e., drawing out the relevant implication or upshot of the talk) as first discussed by Heritage and Watson (1979) but her main focus was on candidate pre-closing formulations. Such types of repeat utterances not only preserve some elements of the preceding talk, displaying a characterization of prior talk for (dis)confirmation but they may also transform or delete other elements of this talk. In this study, the design features of formulations that helped to achieve the link with previous discussions were the use of turn-initial "so" and pro-terms. However, there were not always immediately recognizable elements of prior talk in the case of a transformative gloss of the prior talk and the sequential context determined the utterance to be a formulation.

Barnes's (2007) study demonstrated that in meeting talk, which is characterized by knowledge sharing and an orientation to decision-making, formulation devices played an important role in the display and management of common agreement and shared understanding regarding the outcomes of the talk and in the development of decision-making in a time-limited task-focused environment. Therefore, they had such functions as keeping a track of what is said, bringing forward what is essential or relevant, connecting prior turns or sequences into a coherent whole, developing a sense of cumulative understanding, displaying what is jointly understood, and providing a basis on which to proceed. It was found that the formulation devices were used by the chairpersons for different actions; (1) gist formulations were employed to re-engage the participants and fill a gap in the interaction as well as to display shared understanding, (2) upshot formulations established an action agenda, moving participants toward an outcome while maintaining the topical agenda and thus, they were sequentially both backward and forward looking; and (3) candidate pre-closing formulations helped to manage the task by marking formal acceptance of on-task contributions, closing the business-at-hand, establishing the outcome-so-far of the prior topic talk, facilitating the transition to the next topic and thus, displaying interactional cohesiveness. Being employed as harbingers of agreement on syllabus content, most of the time, formulations required no response. In addition, in the meeting interactions, the

chairpersons not only had the rights to offer a formulation but also had sufficient knowledge about what had been described. Therefore, their use of formulations showed an orientation to their institutional identity in the decision-making process. Overall, the study illustrated that ‘understanding’ in conversation is much more than recognising the literal sense of others’ turns. As Barnes (2007) indicates, formulations are more commonly used in institutional interactions as compared to everyday talk due to their goal-oriented character and the need to explicitly exhibit agreement on points of discussion for task accomplishment.

Adopting an ethnomethodological and conversation analytic perspective, Verkuyten (2000) analyzed a teachers’ meeting where they discussed students’ school reports of a particular class in a secondary school in the Netherlands and provided situated interpretations and explanations of unsatisfactory school marks. More specifically, upon the chairman’s (i.e., the co-ordinator of the program) request for clarification with regard to the unsatisfactory marks in a particular subject, the teachers of the subject in question discussed the progress of their individual students and the potential problems and difficulties in their classes. Additionally, a list containing the students’ names and their marks for each subject was used as a guideline for the discussion. The study examined how the teachers interacted with each other to provide an acceptable explanation for their students’ marks and demonstrated how accounts and explanations were locally managed.

Verkuyten’s (2000) study showed that marking involves teachers’ accountability to colleagues because marking is used not only to evaluate students’ performances but also to provide information for decisions to be taken about students. The teachers used various devices for developing a factual and an acceptable account that emphasized the role of the students in their low marks and that presented a neutral stance for themselves as a way of managing issues of identity. Those interactional devices included “confirmed expectations, extreme case formulations, introducing corroborating witnesses, the deployment of cited others, defining exceptional cases, detailed descriptions and narratives, and the use of different accounts” (Verkuyten, 2000, p.452). It was also found that the teachers referred to the students’ lack of effort and capability as well as their behavioural problems in constructing their explanations for unsatisfactory marks and in creating their arguments. Furthermore, the meeting talk had different sequential consequences. Teachers’ joint construction of accounts and explanations in the meeting had implications for decisions about remedial interventions. That is, decisions for intervention depended on the kinds of

explanations that the teachers gave in interpreting their students' marks. For instance, while 'capacity' account implied that student performances were permanent and difficult to change, 'effort' or 'behaviour' account meant that improvement was possible. Therefore, in the case of the former, it was suggested that those students lacking in ability went down to a lower educational level and in the case of the latter, as a remedial intervention, the teachers agreed to have a corrective talk with their students about their poor performance.

Drawing on membership categorization analysis, Baker (1997) examined a secondary school staff meeting talk concerning the operation of a system of 'ticketing rules' set up for managing student behaviour in a rational and fair manner. This system involved the issuing of tickets to students for their good or bad behaviour in the school based on a guideline produced by the school staff. Certain privileges or punishments in terms of the students' promotion or demotion to a particular level were determined depending on the number of these tickets. However, this system was under constant review and this study examined how the teachers proposed and discussed rule changes in and through interaction in a school staff meeting as part of a moral-organizational work. More specifically, it focused on the accounts of the school staff and how they deployed membership categorization devices in their accounts to persuade, question, and justify as they produced and resolved problems in the system. As a result, the study revealed that the school staff found social-moral order in this system and achieved moral accountability of their actions by engaging in considerable accounting work and by providing category-bound descriptions of people (e.g., students), activities and scenes and referring to principles of justice and fairness in the sequential and consequential production of their accounts. In and through talk, the staff produced "the local rationality and the local morality of their rules and of their rule-making work" (Baker, 1997, p.98). Baker (1997) concluded that through membership categorization devices (i.e., categories of moral actors and of courses of moral action), the underlying school culture was made evident in action in the moment of the meeting.

Hall and Danby (2003) investigated how a group of school teachers and education university academics built and managed a school-university partnership interactionally to collaborate on professional development and research activities to improve Mathematics teaching in a primary school. More specifically, using membership categorization and conversation analysis, the study focused on how they co-constructed the category of expert through meeting talk as they created the business of doing partnership. It was found that both the

teachers and academics constructed themselves as experts and such orientations shaped “what can be said in meeting talk, who gets to speak, and the types of relationships that can be constructed” (Hall & Danby, 2003, p.1). The participants of the meeting described activities that were bound to the membership category of partnership and that showed their understanding of partnerships. As they accounted for their practices, the participants oriented to different categories of persons such as “being a partner” and “being an expert” as a teacher or as a partner and produced descriptions of themselves as experts, academics, teachers, and partners. In one instance, a disjunction was revealed between the principal’s and one teacher’s categorization of “being an expert teacher”, which had consequences for the way the partnership was constructed. While the teacher referred to it as “being an effective teacher”, for the principal it meant “being a teacher leader”. Overall, the study demonstrated how an everyday business of partnership worked in and through a meeting talk and that partnership was a term that was enacted locally without any predefined meanings.

### **2.6.3. Conversation Analytic Studies in Language Testing/Assessment**

In the field of language assessment and testing, CA has recently been used widely to examine the kinds of interactions co-constructed in different L2 oral proficiency test formats such as paired speaking tests (Galaczi, 2008; 2014; Gan et al., 2008; Lazaraton & Davis, 2008; Luk, 2010), group oral tests (Gan, 2010; Nakatsuhara, 2013), language proficiency interviews taking place between testers and testees (Brown, 2003; Kasper & Ross, 2007; Kormos, 1999; Nakatsuhara, 2008; Seedhouse, 2012, 2013; Seedhouse et al., 2014; Young & He, 1998), and guided role-play activities (Kormos, 1999). By using CA, those studies were able to find out (1) the relationship between the nature of interaction and test scores (Galaczi, 2008; Sandlund & Sundqvist, 2011; Seedhouse, 2012; Seedhouse et al., 2014), (2) the effect of certain oral assessment test formats on the interactions and thus, on the students’ display of interactional skills (Gan, 2010; Gan et al., 2008; Kormos, 1999), and (3) how issues such as test-taker extraversion and proficiency levels in group oral tests (Nakatsuhara, 2013), interviewer variation (Brown, 2003; Nakatsuhara, 2008), interlocutor effect in paired oral assessment (Lazaraton & Davis, 2008), task orientations (Sandlund & Sundqvist, 2013), interviewers’ use of multiple questions on the same topic of talk (Kasper & Ross, 2007) and students’ impression management (Luk, 2010) influenced the nature of interactions co-constructed in those oral exams. There are also studies describing L2 oral proficiency interactions as varieties of interaction through CA by investigating their



interactional architecture and the use of certain interactional practices such as repair and then comparing them with ordinary conversation or talk in L2 classroom settings (e.g., Seedhouse, 2013; Young & He, 1998).

Studies of oral language tests showed that CA research can inform the design and implementation of oral assessment tasks as well as rating scales (e.g., Galaczi, 2014; Seedhouse et al., 2014) by explicating the advantages and disadvantages of certain assessment formats and the subtle and complex issues concerning language proficiency assessment (Seedhouse, 2011). They can provide empirical evidence as part of a validity argument and be useful in identifying threats to test validity, remedying problems, and in developing rating scales as part of a data-based approach (Seedhouse, 2013). Therefore, such CA-informed studies dealing with the interactions co-constructed in L2 oral proficiency tests have implications for test design and implementation, test validation, rating scale construction, the development of teaching materials and examiner training (Seedhouse, 2013). With regard to the role of CA in language testing, McNamara (2007) states that “at the micro level, insights from Conversation Analysis have challenged the individualistic focus of current thinking about assessment” and offered “a fully developed social theory of the face-to-face context” (p.131, 133).

In language assessment, CA has also been recently used to investigate the interactional practices that do informal formative assessment in the context of classroom interaction (Can Daşkın & Hatipoğlu, 2019; Heritage & Heritage, 2013). Can Daşkın and Hatipoğlu (2019) focused on the phenomenon named “Reference to a Past Learning Event” emerging in an L2 classroom interaction “when the teacher contingently extends the main instructional activity to focus on what was presented interactionally earlier” (p.1). They demonstrated the relevance of such an interactional phenomenon to informal formative assessment since by referring to a past learning event, the teacher “seeks evidence of student knowledge and understanding and/or acts on the negative evidence already elicited in order to enhance learning and shape instruction” (Can Daşkın & Hatipoğlu, 2019, p.1). On the other hand, Heritage and Heritage (2013) examined a teacher’s questioning practice and suggested that the way teachers construct their questions (i.e., non-threatening pedagogical questioning practices that create a collaborative atmosphere in class) is “a key resource in eliciting students’ current learning status, and for making decisions about next steps in student learning” (p.176), which constitutes the practice of formative assessment. By using CA to

examine a different dimension of language testing and assessment (i.e., IR practice), this study not only brings an interactional perspective to language test construction but also extends the scope of CA and language testing studies. With the accumulation of such micro-level studies of different facets of language testing, a better understanding can be developed with regard to testing practices in educational settings.



## **CHAPTER 3**

### **METHOD OF RESEARCH**

#### **3.0. Introduction**

In this chapter, methodological background, research context, participants, data collection, transcription and analysis procedures are presented and discussed.

#### **3.1. Conversation Analysis as a Research Methodology**

As a methodological approach, CA “aims to describe, analyze and understand talk as a basic and constitutive feature of human social life” (Sidnell, 2010, p.1) and thus, is qualitative in nature (Markee, 2000). More specifically, CA is a methodology for “analyzing talk-in-interaction that seeks to develop empirically based accounts of the observable conversational behaviours of participants that are both minutely detailed and unmotivated by a priori, etic theories of social action” (Markee, 2005, p.355). CA has three broad principles: (1) collecting data, (2) transcribing data, and (3) analyzing data (Wong & Waring, 2010).

In the category of data collection, CA requires the recording of naturally occurring data, which refer to actual instances of talk situated in peoples’ lives and it is not gathered by means of interviews, observation and experiments in laboratories (Hutchby & Wooffitt, 1998). Since CA is concerned with the examination of what speakers actually do as they talk rather than what they think they do as in the case of introspection, actual instances of talk are required as the primary data (Liddicoat, 2007). In this respect, CA has an empirical approach. In CA, video-recording the data is significant because it will provide visual access to the salient features and details of conduct involved in the management of interaction (e.g., gaze direction) (Mondada, 2013). Having the recorded data in hand will also allow the researcher to listen to the original data repeatedly in conjunction with the transcript and help him/her make observations as well as identify previously unnoticed or ignored aspects of interaction (Psathas, 1995; Liddicoat, 2007). Moreover, in video-recording the data, it is important for CA researchers to consider the entire participation framework and set up the

camera(s) accordingly as the aim is to capture the conduct of all participants including the details of both their verbal and embodied behaviour as well as the artifacts, tools and multi-modal resources oriented to by the participants (Mondada, 2013; Ten Have, 2007).

As Psathas (1995) explains, the data in CA is not selected based on some pre-formulated theoretical conceptions, which have been used as the criteria of selection in, for example, laboratory experiments or controlled observations. This is explained by the idea of “unmotivated looking” which refers to discovering what is actually in the data and how participants themselves understand and interpret what they are doing rather than searching for a particular predetermined phenomenon (Psathas, 1995; Wong & Waring, 2010). Thus, the data may be obtained from any available source rather than produced for the purposes of the study. With regard to data collection, obtaining written permission from the participants especially in institutional settings is an important ethical consideration (Psathas, 1995).

As the second principle of CA, transcribing the data is an integral part of CA. As Hutchby and Wooffitt (1998) inform, it is important that the analyst carries out the actual transcription process by repeatedly listening to and watching the recording since this will give the analyst the opportunity to identify some phenomena as the transcription is actually done allowing him/her to get acquainted with the data. In this respect, the act of transcription is, in fact, the beginning of the analysis procedure (Gardner, 2004). In addition, transcription is a process of writing down as many details as possible regarding the key features of talk (e.g., end points of turns, pauses, emphasis, overlaps, cut-offs, silence, hesitation, lengthening of sounds, the relative loudness of talk, laughter token, the relative speed of delivery, marked pitch changes) as well as incorporating information about members’ embodied actions, gestures, and eye gaze behaviours (Markee, 2005; 2007; Seedhouse, 2011). That is, transcripts include details about the dynamics of turn-taking, the characteristics of speech delivery and non-verbal activities (Arminen, 2005). A detailed transcription based on the micro-analytic conventions of CA is necessary to provide the researcher and the reader with sufficient information to understand both what the participants are talking about and how they are talking on a moment-by-moment basis and thus, to be able to describe the orderly practices of social action in interaction (Hepburn & Bolden, 2013; Psathas, 1995). Therefore, a transcription notation system was originally developed by Jefferson (1972) and attempts have been made to improve or extend it. The transcriptions of the data are also expected to be included in reports of CA (Psathas, 1995).

In the data analysis category of CA, a highly empirical approach is taken to analysis and a data-driven approach rather than a data-based approach is followed since interaction patterns are drawn from the data and the researcher does not bring a set of preconceived categories, codes or theoretical assumptions to the data, trying to fit them in the data (Seedhouse, 2004; Walsh, 2011). When carrying out CA, the researcher also does not make assumptions about the contextual details, institutional setting, and the participants' motivations, intentions, ideas, emotions, gender and institutional roles (Psathas, 1995; Seedhouse, 2004). These aspects are discussed only when they are demonstrated in the interaction as participants attend and orient to them, providing evidence for their relevance in the details of talk (Seedhouse, 2004). Besides, in CA, there is not just a focus on the speaker but on all the participants as the talk "is seen as a jointly accomplished activity, with the listener and speaker given equal status as co-constructors of the emerging talk" and "each unit of talk builds upon the prior talk" (Gardner, 2004, p.263).

The first stage in the data analysis procedure is "unmotivated looking" or discovering patterns or phenomena out of an unfocused analysis without basing the database search on any preconceptions and hypotheses (Psathas, 1995; Seedhouse, 2004; Ten Have, 2007). That is, it starts with a case-by-case analysis to uncover the regular procedures displayed by the participants during the normal course of action (Liddicoat, 2007; Wong & Waring, 2010). At this point, it is common for CA researchers to organize data sessions where they share their observations and analyses of the same data segment after viewing and listening to it repeatedly, assisting the researcher in identifying an interesting phenomenon or instances where a well-known phenomenon is used in a specific context (e.g., institutional) and in understanding the features of an interactional event (Kasper & Wagner, 2014; Ten Have, 2007).

The second stage of data analysis is establishing a collection of instances of the phenomenon by inductively searching through the database although a single-case analysis can also be carried out (Seedhouse, 2004). As Markee (2007) explains, a collections-based methodology could be followed in analyzing the data in CA by making collections of individual fragments of talk that appear to be organized in the same way to account for the organization of a practice in the whole corpus. In this respect, data analysis in CA begins with observation and unmotivated examination of the data, the goal of which is to notice recurrent details of talk and identify a certain phenomenon that can then allow for a collection of comparable

instances to be built and a precise description of the phenomenon (Kasper & Wagner, 2014; Sidnell, 2013; Ten Have, 2007). The next stage is to establish regularities and patterns in relation to the way the phenomenon is interactionally realized and to show that the participants orient to them as normative organizations of action (Seedhouse, 2004). This is followed by the micro-analytic analysis of single instances of the phenomenon so that the organization of the uncovered patterns can be explicated (Seedhouse, 2004). At this stage, besides the regular procedures displayed by the participants, the deviant cases are considered as well since they are also seen to be an indication of orderliness in talk and are to account for the apparent exceptions, contributing further to the understanding of the identified phenomenon and serving to demonstrate the normativity of practices (Liddicoat, 2007; Markee, 2007; Seedhouse, 2004; Wong & Waring, 2010). In this way, the characteristics of interaction in a particular variety of institutional discourse as a whole and in its sub-varieties can be described for a complete overall investigation of the corpus (Seedhouse, 2011; Ten Have, 2007). Finally, a more generalized account is produced with respect to “how the phenomenon relates to the broader matrix of interaction” (Seedhouse, 2004, p.39).

Since CA “aims to explicate how members orient (that is, observably pay attention) to certain behavioural practices as they co-construct talk-in-interaction in real time” (Markee, 2005, p.355), in analyzing the data, conversation analysts focus on the basic interactional practices of (1) turn-taking, (2) sequence organization (e.g., adjacency pairs, topic management), (3) repair, and (4) turn design, which significantly organize interaction (Wong & Waring, 2010). In addition to these primary units of analysis, the dimensions of overall structural organization (e.g., openings and closings), lexical choice and forms of interactional asymmetry are proposed for the analysis of institutional talk (Drew & Heritage, 1992) (See Section 2.6.1). CA explains not only how utterances are constructed grammatically and lexically and how turns are designed but also how turns are taken orderly, how they are sequenced into sets of actions as adjacency pairs and more extended sequences, providing a full account of the mechanics of talk (Gardner, 2004). During the CA analysis of data, Seedhouse (2011) suggests asking the question “Why that, in that way, right now?”, which reflects the view of interaction as action (i.e., why that?) expressed via linguistic forms (i.e., in that way) in a developing sequence (i.e., right now). The goal in doing such a micro-analytic analysis is to examine “the fine detail of talk, of the underlying structures that members of the social group draw upon to constitute their social world” and to describe what is hearable and observable in the data within the context of the talk (Gardner, 2004, p.267).

In short, Ten Have (2007) summarizes what makes CA different from other approaches. According to him, (1) CA reveals details of human interaction that are lost in other practices by working with transcriptions, (2) CA favours naturally occurring data rather than experimental data, (3) CA views interaction as organizational, procedural, and as a collectively organized event and aims to explain how people act the way they do rather than why and (4) CA studies verbal language as it is used interactionally rather than written language as in many traditional forms. According to Lazaraton and Taylor (2007), CA, which has recently been employed in language testing especially in studies of oral assessment, is an inductive method for finding recurring patterns of talk-in-interaction in transcribed authentic, recorded data by using the turn as the unit of analysis. In this approach, the research questions are not pre-stated and the data is treated with an unmotivated looking. The data is also not coded or quantified and the ethnographic and demographic elements of the context and participants are disregarded unless otherwise demonstrated in the data (Lazaraton & Taylor, 2007).

### **3.2. Research Context**

In Turkey, since the introduction of foreign language instruction into the Turkish education system and due to the tremendous need to use English internationally in the scientific and technical fields of industrialized countries, there has been an increasing demand for both state and private universities to offer English-medium or mixed Turkish-English medium education at universities. The need for an English education at universities has been satisfied through preparatory schools which give students a full one year of English education with at least 25 hours of English per week. Those schools enable students to have a proficient knowledge of English by developing their abilities to comprehend what they read and listen, express themselves in oral and written contexts, and communicate in cultural and social life in English so that they can follow their courses in their departments effectively.

As for English language assessment practices in preparatory schools in Turkish universities, at the beginning of the academic year, students take a placement test, which is generally prepared by the institution itself unless they have a score from an internationally accredited English exam such as TOEFL or IELTS or a score obtained from the language proficiency exam administered by Higher Education Council Student Selection and Placement Centre. Throughout the English language program at the preparatory school, students take quizzes,

progress and achievement tests, which are locally determined, and depending on their level of success, they take two proficiency exams- one at the end of the first semester and another at the end of the year when they complete the whole program. In the majority of preparatory schools in Turkey, which tend to have a large population of teachers and learners, tests are locally developed by testing offices rather than individual teachers to ensure standardization and fairness. Since there are many teachers working in those institutions and many groups of learners, teachers constructing their own tests and testing learners consistently may become a problem due to individual differences among teachers, who might have different testing practices. Thus, teachers already working as instructors in the preparatory school are usually assigned as testers at the testing office of the school and each tester is responsible for preparing the tests of a particular group of students at a certain level of proficiency. Testers are expected to develop a representative sample of what the students have been taught in the language program and to consider the teaching practices and learning activities while constructing the tests to make sure that the tests are in line with the course content and the objectives of the curriculum. Generally, instructors rotate through the testing office, and all the instructors do the checking and scoring of their students' tests. Therefore, the majority of the instructors somehow get involved in the test construction practice and are assigned as testers at some point in their lives at the institution. In addition, in most of the cases, the tester of each language proficiency group works in close cooperation with the academic coordinator (i.e., Curriculum office) and the teachers of that group, who review the achievement tests, evaluating and making comments about their effectiveness. They may offer suggestions for modifications needed and actively cooperate with the testers in revising and editing the tests. The practice of item reviewing (IR) and the selection of reviewers from among the current teachers helps to ensure that the tests are parallel with teaching and to avoid any mismatches between the tests and instructional events. Thus, most of the teachers are somehow involved in the testing process and they can always give feedback to the testers and coordinators about the tests.

In a study conducted in the preparatory schools of 14 state universities in Turkey, Oğuz (2003) found that the testing instruments were mainly testing-office made and instructors had limited authority over the choice of assessment instruments. She also revealed that the preparatory schools often used traditional testing instruments consisting of, for example, multiple-choice, fill-in-the-blanks, matching, short-answer or true-false items and that the instructors found the testing practices satisfactory as they directly assessed what was taught.



However, a study conducted on the state of English in Higher Education in Turkey by British Council (2015) showed that in the preparatory schools in Turkey, the quality of in-house tests were quite variable and while some universities had invested heavily in training members of their testing units, others seemed to be much less professional.

The data for this study was also obtained from a preparatory school at a state university in central Turkey and thus, the research context described in this study can be considered to be representative of the testing practices at prep schools in Turkish universities and in EFL contexts at tertiary level in general. As in the case of all preparatory schools in Turkish universities, the aim of this school is to equip students who have just entered university with the English language knowledge and skills that are necessary to move to and ‘survive’ in their departments since the medium of instruction in the majority of departments at this university is English.

The curriculum of the prep school where the data for this study was obtained was prepared locally by the EFL teachers who were selected to work at the curriculum office. The school had a year-long English language program which was structured around the objectives identified based on the Common European Framework of Reference. The program had a syllabus-based approach and consisted of fundamental courses as well as two skill courses named “listening and speaking” and “reading and writing”. The fundamental courses followed a textbook selected by the school (i.e., New English File) which included integrated English lessons with a main focus on grammar topics and vocabulary items, and which constituted the majority of the program. On the other hand, in the two skill courses, certain skill-based books and mainly supplementary materials prepared by the curriculum office were used.

The prep school had an exam-oriented system and the students were subject to a number of progress and level achievement tests during the course of study. At the time of data collection, a year-long education in the English preparatory school was divided into four quarters, each of which took seven weeks. In each quarter, students, who were placed at a particular level of English according to Common European Framework (e.g., A2, B1), were required to take five written exams referred to as quizzes, a level achievement test and a speaking exam, which were prepared by the testing unit supervisors at the prep school. The written exams tested not only grammar and vocabulary but also reading, listening and

writing skills. One of the quizzes consisted of a story exam. Students at each level were assigned a story book to be read independently outside class and were tested on their comprehension of the story in a quiz. Their success in those progress tests and the level achievement test, which they took at the end of the quarter, determined whether they were eligible to move onto the next level of proficiency in the following quarter of the program. In addition, if the students fulfilled the requirements during the course of study, they took an English proficiency exam either at the end of the first semester and/or at the end of the academic year to determine whether they had met the required exit level of English proficiency and thus, their readiness to start their departments.

As in many university prep schools in Turkey, at the prep school where the data for this study was obtained, the testers in the testing office prepared the exams for the language proficiency level they were responsible for and the students at the same level all took the same exams for purposes of standardization. The testing unit supervisors were selected from among the instructors already working at the prep school and after working at the testing office for some time, they could take it in turns. Instructors who were responsible for test construction had quite a few hours of teaching and their main responsibility was to construct tests for a specific group of learners, administer it and make decisions about scoring and the interpretation of results.

#### *The practice of item reviewing in the prep school's testing office*

In the preparatory school where the data for this study was collected, there is a certain procedure that the testing unit supervisors have to follow in preparing the exams. The most important part of this process is the reviewing of exam papers by the teachers and this study focused on the teacher and tester interactions that took place in those IR meetings at the testing office. After the testers prepare an exam individually in line with the materials covered in class, they select a couple of teachers (at least one teacher) and invite them to do the reviewing of the exam developed and to provide feedback especially considering their classroom teaching and learning context, co-constructing the exams together. Those IR sessions are an important part of the teachers' practices at this institution as they shape the exams to a great extent and help to foresee future problems that might negatively affect the students and the stakeholders including the teachers and testers. Therefore, an instructor working at this institution is responsible not only for teaching but also for collaborative

works such as reviewing of exams and materials. However, every teacher might have a different perception or conceptualization of IR and thus, the effectiveness of every IR session might change depending on the reviewer's interaction. Therefore, there seems to be a need for a framework of how to conduct language test IR systematically to achieve standardization and obtain the maximum benefit from these sessions on the part of the testing unit supervisors as well as the teachers and learners.

As for the procedure of IR practice in the testing offices, the selected teachers, who are generally experienced in both teaching and testing and who are reflective and can provide constructive feedback, go to the testing office at the arranged time and are given a copy of the exam paper. At first, they answer the questions as a student would and critically think about each one of them on their own. They write down their notes next to the questions. Then the reviewing of the exam starts and the teachers and testers together begin to review the exam interactionally. At this point, although the teachers (i.e., the reviewers) are not given a particular checklist or a guideline to follow and are not involved in a formal procedure in reviewing the items, as part of the IR tradition at the institution, they already know what is expected of them for the IR to be effective and act accordingly in the IR sessions. For instance, they are generally expected to give feedback on;

- the relevancy of the testing points and tasks by reflecting on their teaching practices and the materials covered in class,
- the points that would not be worthwhile to test by considering what was emphasized in class and what was of little importance,
- the validity of the items (i.e., do they test what they are supposed to test?),
- the difficulty level of the items by considering their students' proficiency level,
- the appropriacy of the items in terms of the testing principles especially if they are experienced and trained in testing,
- the accuracy and appropriacy of the language used in the items and on whether there are any ambiguities and any parts open to different interpretations.

The reviewers are also expected not only to state and explain the problem but also to provide suggestions to improve the problematic items and to co-construct them collaboratively. On the other hand, if the invited teachers are 'novice' reviewers, when they are given the copy of the exam paper, the testers verbally tell them what they are expected to do in reviewing the exam. For instance, they tell those teachers to evaluate the items from the teaching or students' perspectives and to think of some alternatives for the problematic items. As the

reviewing takes place, the testers take notes and after the session, make the necessary revisions and corrections ready for photocopying and administration. If too many changes are made, the exam goes through another reviewing done by a different group of teachers. The IR sessions generally take 2-3 hours depending on the length of the exam and are carried out in Turkish.

Overall, the teachers who review the tests act as a bridge between testing and teaching and examine the tests from the students' perspective in the light of their classroom teaching and learning context, instructional events, practical concerns and the program content as well as their testing knowledge and experience so that the prepared tests can be improved in terms of their validity, reliability and fairness. The conjointly construction of tests by teachers and testers together helps to dig out the aspects that could go unnoticed in the first draft of the tests as they together talk about the strengths and weaknesses of the test items, negotiate, come to terms, take certain decisions and eventually formulate the final copy. In this respect, in the IR sessions in test design, the reviewers do much more than a key check as IR practice is an important part of the test validation procedure, aiming to create valid tests and to prevent future problems that would possibly emerge.

### **3.3. Participants**

In total, 7 testers (i.e., teachers assigned as testers in the testing office of the school) and 25 teachers (i.e., reviewers) participated in the study. As Table 1 shows, of the teachers, who did the reviewing of the exams, all were female except one teacher and their ages ranged between 29 and 52. In addition, they were all experienced teachers having a teaching experience of 7 to 30 years. The teachers also had 17-25 hours of teaching a week. Except 7 teachers, they had all worked at the testing office in the past with their testing experience ranging between 1 and 9 years. Therefore, it can be said that most of the reviewers were experienced in both teaching and testing. As for their academic career, while 3 teachers were still doing their masters, 14 of them had already completed it in fields such as English Language Teaching, Educational Administration, English Language and Literature, and Linguistics. 6 teachers further continued with their doctoral studies and while 3 of them had already completed it, the doctoral studies of 3 teachers were still in progress. Of the 25 teachers, only 8 teachers had not continued with a graduate study.

**Table 1. Teachers' Demographic Information**

| <b>Teachers</b> | <b>Age</b> | <b>Gender</b> | <b>B.A.</b>                     | <b>M.A. / P.h.D</b>                                                         | <b>Teaching Experience</b> | <b>Testing Experience at Testing Office</b> |
|-----------------|------------|---------------|---------------------------------|-----------------------------------------------------------------------------|----------------------------|---------------------------------------------|
| BA              | 47         | Female        | English Language and Literature | M.A & P.h.D in English Language and Literature                              | 24 years                   | None                                        |
| DE              | 47         | Female        | English Language Teaching       | M.A in English Language Teaching                                            | 21 years                   | 8 years                                     |
| YE              | 40         | Female        | English Language and Literature | None                                                                        | 12 years                   | 4 years                                     |
| ES              | 39         | Female        | English Linguistics             | M.A in English Language Teaching                                            | 16 years                   | 9,5 years                                   |
| GA              | 34         | Female        | English Linguistics             | None                                                                        | 11 years                   | None                                        |
| CI              | 29         | Female        | English Language and Literature | M.A in English Language and Literature                                      | 7 years                    | 1,5 years                                   |
| AD              | 48         | Female        | English Language and Literature | M.A in English Language and Literature                                      | 25 years                   | None                                        |
| GE              | 48         | Female        | English Language and Literature | M.A in English Language Teaching (in progress)                              | 24 years                   | None                                        |
| SE              | 38         | Female        | American Culture and Literature | M.A in TESOL, P.h.D in Teaching Turkish as a Foreign Language (in progress) | 14 years                   | 2 years                                     |

**Table 1 (continued)**

|    |    |        |                                 |                                                                                     |          |         |
|----|----|--------|---------------------------------|-------------------------------------------------------------------------------------|----------|---------|
| DU | 44 | Female | English Language Teaching       | M.A in Linguistics                                                                  | 20 years | 7 years |
| SU | 39 | Female | English Language Teaching       | M.A in English Language Teaching, P.h.D in Educational Administration (in progress) | 18 years | 2 years |
| OZ | 46 | Female | English Language Teaching       | M.A in TEFL                                                                         | 24 years | 3 years |
| AV | 45 | Female | English Language Teaching       | M.A & P.h.D in English Language and Literature                                      | 20 years | None    |
| MI | 46 | Female | English Language Teaching       | M.A in English Language Teaching (in progress)                                      | 22 years | 1 year  |
| AS | 49 | Female | Translation and Interpretation  | None                                                                                | 28 years | 2 years |
| TA | 48 | Female | English Language and Literature | M.A in British Cultural Studies                                                     | 24 years | None    |
| AB | 48 | Female | English Linguistics             | None                                                                                | 24 years | 2 years |
| YU | 29 | Male   | English Language and Literature | None                                                                                | 7 years  | 1 year  |
| SO | 39 | Female | English Language Teaching       | M.A in Curriculum and Instruction, P.h.D in Educational Administration              | 17 years | None    |
| LE | 43 | Female | English Language Teaching       | None                                                                                | 21 years | 2 years |

**Table 1 (continued)**

|    |    |        |                                 |                                                                                     |          |         |
|----|----|--------|---------------------------------|-------------------------------------------------------------------------------------|----------|---------|
| FU | 40 | Female | English Linguistics             | M.A in English Language Teaching, P.h.D in Educational Administration (in progress) | 16 years | 3 years |
| CA | 46 | Female | English Language and Literature | None                                                                                | 24 years | 3 years |
| KU | 37 | Female | English Language Teaching       | M.A in Educational Administration (in progress)                                     | 15 years | 2 years |
| AK | 52 | Female | English Language Teaching       | M.A. in English Language Teaching                                                   | 30 years | 1 year  |
| SI | 38 | Female | English Language Teaching       | None                                                                                | 16 years | 2 years |

The teachers who participated in this study were all experienced test reviewers at the institution and they were commonly invited to do the reviewing of the exams. According to them, this was due to their experience in the testing office and/or their positive approach to IR as they were cooperative, careful and analytic with a keen eye for detail. They also tried to come up with alternatives or solutions for the problematic items and were willing to co-construct them collaboratively with the testers. To illustrate, one of the teachers said, “I think there is an idea that some teachers are good reviewers because they give comprehensive feedback and are able to see the details that may be problematic” (CI).

Regarding the 7 testers, who prepared the exams for the level of proficiency they were responsible for, as Table 2 shows, all were female and their ages ranged between 24 and 36. Thus, the testers were relatively younger than the reviewers in general. In terms of the teaching experience, only one tester was a novice teacher with only one year of teaching and the rest of the testers were experienced in teaching with 5 to 14 years of teaching. However, in general, the testers were less experienced in teaching than the teachers who participated in this study as reviewers. The testers also had only 4 hours of teaching a week while working

at the testing office. In addition, they had been working at the testing office for some time, ranging from 10 months to 3 years. As for their academic career, all of the testers except two continued with their graduate studies in fields such as English Language Teaching and Educational Administration.

**Table 2. Testers' Demographic Information**

| Testers | Age | Gender | B.A.                            | M.A. / P.h.D                                           | Teaching Experience | Testing Experience at Testing Office |
|---------|-----|--------|---------------------------------|--------------------------------------------------------|---------------------|--------------------------------------|
| HA      | 36  | Female | English Language Teaching       | M.A in Guidance and Psychological Counselling          | 14 years            | 2,5 years                            |
| AY      | 24  | Female | English Language Teaching       | M.A in English Language Teaching (in progress)         | 1 year              | 10 months                            |
| SA      | 32  | Female | Translation & Interpreting      | M.A in Educational Administration                      | 10 years            | 3 years                              |
| ZU      | 30  | Female | English Language and Literature | None                                                   | 8 years             | 2 years                              |
| AR      | 35  | Female | English Language Teaching       | M.A & P.h.D in English Language Teaching (in progress) | 13 years            | 1 year                               |
| ME      | 28  | Female | English Language Teaching       | M.A in English Language Teaching (in progress)         | 5 years             | 2,5 years                            |
| EK      | 35  | Female | English Language and Literature | None                                                   | 13 years            | 2,5 years                            |



When both the teachers and testers were asked of their opinion on their testing experience at the testing office, they all indicated that it was stressful, tiring, difficult and challenging but enjoyable and creative at the same time as new things were continuously being produced. One of the testers, for instance, stated that working at the testing office “is stressful because you produce something in a limited time and the final work does not accept any fault. Trying to develop faultless tests can make you feel under pressure” (HA). They also stated that it was the practice of IR, teachers’ or experienced colleagues’ feedback and their own experience and mistakes in constructing tests that helped them most to improve their testing skills at the testing office. They benefited from the testing experience to a great extent in that as the testers pointed out, it helped them to focus more on details and become more creative and critical in terms of the materials and exams. In relation to this, the teachers who had previously worked at the testing office for some time mainly indicated that their testing experience affected their teaching positively, improving their teaching and their ability to design materials as they gained the ability to consider both sides, that is both teaching and testing. As one of the teachers said, “I benefited a lot because I learned how to prepare an exam and when I continued teaching later, it helped me think of how what I teach will be tested” (CI). The testing experience was also said to improve teachers’ personal skills such as stress management and problem solving skills, their human relations and ability to work as a team.

Regarding the teachers’ views on the practice of IR in language test construction, it was believed that it helped to evaluate the test items from different viewpoints and gave testers the opportunity to talk to teachers about the level of the students, the instructional events and how they were handled in class so that it could be ensured that the exam was in line with the curriculum and the teaching practice at the institution. A careful review of the test items was also said to reduce the risk of errors as the testers who actually prepared the exams may not be able to notice their own mistakes, in which case a ‘second eye’ might be required. Therefore, it was thought that IR practice made language test construction process quite productive and fruitful and the test itself fair for the students. Some of the teachers’ comments on IR are given below.

Colleagues at the testing units sometimes fail to notice the shortcomings / mistakes in the exams they prepare. Therefore, careful item reviewing done by trusted instructors with dedication to the curriculum and significant work experience is essential. It definitely reduces the risk of errors with embarrassing consequences and it leads to an increase in the quality of exams. (AD)

I think it is very important because ones who prepared the exams cannot see their own mistakes. Sometimes simple spelling, grammar mistakes may be there and no matter how many times one looks, he/she cannot or at least may not be able to see them on his/her own. Secondly, other teachers have different perspectives and they may be able to look at the exam in a way which the person who prepared it cannot. (CI)

I think it is one of the best things that is done in order to make things fair for the students. Before there is an actual problem, the item reviewing practices prevent it. The teachers can foresee the problems as they are the ones who actually teach and help prepare an effective exam. (SE)

Different points of view add positive contributions to a test. (OZ)

Additionally, when the teachers were asked what made an IR session more effective, they listed the following;

- a good interaction between the reviewer and the tester,
- working as a team with at least two reviewers,
- discussing the items all together,
- being open to different views and criticisms about the written product,
- listening to one another,
- trying to understand each other,
- finding a solution that everyone can agree on,
- trying to match the course objectives with the testing points,
- checking the level of the test, and
- putting oneself in students' shoes.

One teacher pointed out that both the reviewers and the testers should get theoretical and practical training and be experienced in teaching at the institution so that they can understand the level and needs of the students. Another teacher stated that finding the right reviewers was important for the effectiveness of the IR practice. With regard to some of those points, one of the teachers OZ stated that “reviewers should focus on the work, not the person who has prepared the test. They should consider appropriacy as well as difficulty. Testers should not be defensive. They need to approach the task professionally. They need to know that it is not them that are criticized. Both parties should accept this process is for the good of student”.

The testers also had similar views on the practice of IR. Like the teachers, they thought that it was the most important and effective part of test construction as it brought different viewpoints, allowing them to see what they could not previously see and helped to evaluate

the test critically, leading to necessary revisions in the test items. In their comments, they focused more on the quality of the IR and noted that it should be done attentively. According to them, test IR became an important practice if it was carried out successfully and that the reviewers determined the effectiveness of the process as some teachers may not be willing to review the exam in detail. Furthermore, the testers thought that for an IR session to be more effective, all of the reviewers should meet at the same time; discuss together the test items one by one, taking into consideration all the parameters of the exam; clearly explain why a question would not work; and rewrite or revise the problematic items by offering some alternatives. This was illustrated in ME's comment who said that "when teachers explain why a question does not work for students in a detailed way and give some alternatives, I develop my horizon. Discussing the questions and finding alternatives together make the process effective". One of the testers also pointed out that the tester should be ready and fully prepared for IR by, for instance, putting the exam material in a real exam format. Some teachers indicated that the teachers (i.e., reviewers) should aim to provide constructive feedback rather than just criticize and they need to be well-qualified, having a good command of the subject matter and the course objectives.

When choosing their reviewers, the testers had some criteria. One of the main criteria was the teachers' experience in both teaching and testing and their dedication to their job and the other was their degree of cooperation and their approach to IR. Generally teachers with a critical eye and those who were eager to give suggestions and helpful feedback and who were more likely to share their ideas were preferred for an IR. One tester stated that she usually chose her close friends as a reviewer as she could get feedback from them in a relaxed atmosphere and another said that personal relations mattered in choosing her reviewers.

### **3.4. Data Collection Procedures**

A corpus of EFL teachers' IR interactions for syllabus-based language tests in an English Preparatory School in a Turkish higher education setting constitutes the data for this study. 25 hours of video and audio recordings of syllabus-based language test IR sessions taking place among maximum three classroom language teachers and a tester were collected at a state university in Turkey. The interactions most often consisted of teacher-tester dyads or triads including two teachers and the tester. The details about the IR data are given in Table

3. This data is naturally occurring, involving no interference and any kind of changes in the IR practice and was recorded as it normally took place without the researcher's presence. Before the video-recordings, the participants were informed about the purpose of the study and their written permission was obtained (See Appendix E for the consent form). They were also assured that the data would only be used for scientific purposes, their identities would be kept confidential by, for example, using pseudonyms, and the participation in the study was on a voluntary basis.

During the process of collecting the data, since the object of the video-recordings was only small groups (often dyads and triads) working together on a task in a compact way around a table, one camera was set up and two audio-recorders were placed on two ends of the table around which the participants were seated so that anything missed by the camera could be captured. Considering the task-oriented nature of the recorded interactions, they included the use of documents (i.e., test materials, textbooks) and thus, the camera was positioned in such a way that it recorded the entire participation framework, capturing not only the verbal and embodied conduct of all the participants but also the artifacts that they oriented to by writing, reading or looking (Mondada, 2013; Ten Have, 2007). In this way, the details of the language test IR interactions could be documented as a multi-modal activity. With respect to the physical arrangement of the tables in the testing offices where the IR practice was conducted, they were rectangular-shaped tables positioned up against a wall and the participants gathered around one side of the table, leaving one of its other sides unseated, which provided space for the camera and enabled it to record the relevant participation framework including all the participants.

Before the recordings actually began, the researcher went into the testing offices and set up the devices well in advance as the reviewers slowly appeared and were given a copy of the test for them to take individually as a student would prior to the IR session. In this way, the participants were given the chance to get accustomed to the presence of the camera as they concentrated on the task at hand and thus, an attempt was made to make the activity of recording as unobtrusive as possible and to overcome observer's paradox (Allwright & Bailey, 1991). After the video-recording, both the teachers and the testers were given a background questionnaire (See Appendix A) to get their demographic information and to learn about their teaching and testing experiences as well as their views on the IR practice in their institution so that a thick description could be provided about the participants and the

context where this study was carried out (See Section 3.3). In addition, following the administration of the reviewed test, a copy of the final version of the test was obtained from the testers to be able to get an idea about the test task and refer to it while transcribing the data when needed. For confidentiality purposes, a copy of the reviewed draft of the tests could not be obtained before the IR sessions.

The reviewed tests in the data followed a syllabus-content approach and aimed to assess the progress and achievement of the learners on the material covered in class throughout a quarter, which made a period of seven weeks, to determine if they were ready to move to the next level of proficiency. During a quarter, students' progress was assessed by means of 5 quizzes, which consisted of either one of the use of English, listening and reading sections, and each quiz was worth of ten points. One of the quizzes involved a story exam, which tested students' comprehension of a story that they were assigned by the curriculum office to be read independently outside class. The level achievement test, which was worth of 100 points, was administered at the end of the quarter to assess students' overall success of the program at a certain English level and to determine whether they had achieved the objectives of this level of proficiency. This test included the use of English, reading, listening and writing sections. The data in this study consists of the language test IR interactions for the quizzes and the level achievement tests that the students took during a quarter.

As Table 3 shows, in total, 26 language test IR sessions were recorded in four different quarters throughout the academic year via convenience and opportunistic sampling (Dörnyei, 2007). The recordings included the review of tests at various proficiency levels including A2 (N=1), B1 (N=6), B1 plus (N=10), B2 (N=7), B2 plus (N=1) and C1 (N=1). Five of the tests in the data were reviewed more than once by different groups of reviewers at different times. As Table 3 shows, the tests in the IR sessions 3 and 4; 7 and 8; 13 and 14; 17, 18 and 19; and those in the sessions 20 and 21 were reviewed by different groups of reviewers on successive days or at different times during a day. Except the two tests reviewed in sessions 3 and 4 as well as in 17, 18 and 19, it was the revised versions of the other three already reviewed tests that were further evaluated and revised in their second IR sessions.

**Table 3. Details about the Language Test IR Data**

| <b>No</b> | <b>Duration<br/>(Minutes)</b> | <b>Level</b> | <b>Test Type</b>                                  | <b>Task Type</b>                                                                                               | <b>Participants</b>           |
|-----------|-------------------------------|--------------|---------------------------------------------------|----------------------------------------------------------------------------------------------------------------|-------------------------------|
| <b>1</b>  | 34                            | B1           | Quiz 2 / Use<br>of English                        | Fill-in-the-blanks (Tense<br>passage)<br>Binary-choice                                                         | SA (Tester)<br>CI<br>YE       |
| <b>2</b>  | 24                            | A2           | Quiz 3 /<br>While-<br>Listening                   | Fill-in-the-blanks (Chart<br>completion)                                                                       | HA (Tester)<br>SU<br>TA       |
| <b>3</b>  | 63                            | B1+          | Quiz 3 / Use<br>of English                        | Multiple-choice cloze test<br>Fill-in-the-blanks (Dialogue<br>completion)                                      | ME (Tester)<br>MI<br>AV<br>AD |
| <b>4</b>  | 21                            | B1+          | Quiz 3 / Use<br>of English                        | Multiple-choice cloze test                                                                                     | ME (Tester)<br>AD             |
| <b>5</b>  | 11                            | B2           | Quiz 5 / Story<br>Exam                            | True/False<br>Matching<br>Short-answer                                                                         | EK (Tester)<br>AB             |
| <b>6</b>  | 15                            | B2           | Quiz 2 / Note-<br>Taking<br>Listening             | True/False                                                                                                     | EK (Tester)<br>FU             |
| <b>7</b>  | 119                           | B1+          | Quiz 3 / Use<br>of English                        | Multiple-choice cloze test<br>Rewrite (Sentence<br>completion)                                                 | ME (Tester)<br>AD<br>CA       |
| <b>8</b>  | 85                            | B1+          | Quiz 3 / Use<br>of English                        | Multiple-choice cloze test<br>Rewrite (Sentence<br>completion)                                                 | ME (Tester)<br>ES<br>LE<br>KU |
| <b>9</b>  | 38                            | B1           | Quiz 3 / Note-<br>Taking &<br>While-<br>Listening | <i>Note-taking:</i><br>Fill-in-the-blanks (Outline<br>completion)<br><br><i>While-listening:</i><br>True/False | SA (Tester)<br>DU<br>GE       |
| <b>10</b> | 34                            | C1           | Quiz 5 / Note-<br>Taking<br>Listening             | Multiple-choice<br>Fill-in-the-blanks                                                                          | EK (Tester)<br>AS             |
| <b>11</b> | 52                            | B1+          | Quiz 5 / Story<br>Exam                            | Short-answer<br>True/False<br>Matching                                                                         | ME (Tester)<br>YU             |

**Table 3 (continued)**

|           |     |     |                                                                       |                                                                                                                                                                                                                                                                                        |                               |
|-----------|-----|-----|-----------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------|
| <b>12</b> | 128 | B2+ | Level<br>Achievement<br>Test                                          | <i>Use of English:</i><br>Multiple-choice cloze test<br>Rewrite (Sentence completion)<br>Fill-in-the-blanks (Word formation)<br><br><i>Reading:</i><br>True/False/Not given<br>Short-answer (Referencing)<br>Multiple-choice<br><br><i>Listening:</i><br>Multiple-choice<br>True/False | ZU (Tester)<br>AK             |
| <b>13</b> | 126 | B1+ | Level<br>Achievement<br>Test (Use of<br>English and<br>Reading parts) | <i>Use of English:</i><br>Multiple-choice cloze test<br>Multiple-choice rewrite and dialogue<br><br><i>Reading:</i><br>Multiple-choice                                                                                                                                                 | ME (Tester)<br>SO<br>CA<br>AD |
| <b>14</b> | 63  | B1+ | Level<br>Achievement<br>Test (Reading<br>part)                        | Multiple-choice                                                                                                                                                                                                                                                                        | ME (Tester)<br>AD             |
| <b>15</b> | 117 | B1  | Level<br>Achievement<br>Test                                          | <i>Use of English:</i><br>Multiple-choice cloze test<br>Multiple-choice rewrite and dialogue<br><br><i>Reading &amp; Listening:</i><br>Multiple-choice                                                                                                                                 | SA (Tester)<br>DU<br>GE       |
| <b>16</b> | 67  | B2  | Level<br>Achievement<br>Test                                          | <i>Use of English:</i><br>Multiple-choice cloze test<br>Rewrite (Sentence completion)<br>Fill-in-the-blanks (Vocabulary passage)<br><br><i>Reading:</i><br>Multiple-choice                                                                                                             | EK (Tester)<br>BA<br>DE       |

**Table 3 (continued)**

|           |    |     |                                                                        |                                                                                                                                                                                                                                                                                       |                         |
|-----------|----|-----|------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|
| <b>17</b> | 33 | B2  | Level<br>Achievement<br>Test<br>(Use of<br>English &<br>Reading Parts) | <i>Use of English:</i><br>Multiple-choice cloze test<br><br><i>Reading:</i><br>Matching (Heading)<br>Multiple-choice<br>True/False/Not given<br>Fill-in-the-blanks<br>(Summary completion)<br>Short-answer (Referencing)                                                              | AR (Tester)<br>SE       |
| <b>18</b> | 26 | B2  | Level<br>Achievement<br>Test (Reading<br>part)                         | True/False/Not given<br>Fill-in-the-blanks<br>(Summary completion)                                                                                                                                                                                                                    | AR (Tester)<br>AK       |
| <b>19</b> | 57 | B2  | Level<br>Achievement<br>Test                                           | <i>Use of English:</i><br>Multiple-choice cloze test<br><br><i>Reading:</i><br>Matching (Heading)<br>Multiple-choice<br>True/False/Not given<br>Fill-in-the-blanks<br>(Summary completion)<br>Short-answer (Referencing)<br><br><i>Listening:</i><br>True/False<br>Fill-in-the-blanks | AR (Tester)<br>SE<br>AK |
| <b>20</b> | 33 | B1  | Quiz 3 / Use<br>of English                                             | Binary-choice test<br>Multiple-choice dialogue<br>and rewrite                                                                                                                                                                                                                         | AY (Tester)<br>CI       |
| <b>21</b> | 44 | B1  | Quiz 3 / Use<br>of English                                             | Binary-choice test<br>Multiple-choice dialogue<br>and rewrite                                                                                                                                                                                                                         | AY (Tester)<br>SI       |
| <b>22</b> | 60 | B1+ | Quiz 3 / Use<br>of English                                             | Multiple-choice cloze test<br>Rewrite (Sentence<br>completion)                                                                                                                                                                                                                        | ME (Tester)<br>BA<br>DE |
| <b>23</b> | 52 | B1+ | Quiz 5 / Story<br>Exam                                                 | Short-answer<br>True/False<br>Matching                                                                                                                                                                                                                                                | ME (Tester)<br>OZ       |
| <b>24</b> | 58 | B1  | Quiz 5 / Story<br>Exam                                                 | True/False<br>Short-answer<br>Matching                                                                                                                                                                                                                                                | AY (Tester)<br>YU       |
| <b>25</b> | 81 | B1+ | Level<br>Achievement<br>Test (Reading<br>& Listening<br>Part)          | <i>Reading:</i><br>Multiple-choice<br><br><i>Listening:</i><br>Multiple-choice<br>True/False                                                                                                                                                                                          | ME (Tester)<br>GA       |



**Table 3 (continued)**

|                                         |    |    |                                                       |                            |                   |
|-----------------------------------------|----|----|-------------------------------------------------------|----------------------------|-------------------|
| 26                                      | 51 | B2 | Level<br>Achievement<br>Test (Use of<br>English Part) | Multiple-choice cloze test | AR (Tester)<br>MI |
| <b>TOTAL: 1492 minutes<br/>25 hours</b> |    |    |                                                       |                            |                   |

Of the tests reviewed in the data, there were 16 quizzes and 10 level achievement tests. 8 of the quizzes consisted of the use of English section including both grammar and vocabulary; 4 of them tested students' listening skill and the remaining 4 quizzes tested students' comprehension of a story book that they had read independently. Of the 10 level achievement tests, which included the use of English, reading, listening and writing sections, the reviewing of 7 use of English, 9 reading and 5 listening sections was recorded. Thus, the most frequently reviewed section in the data was the use of English section (N=15) followed by reading comprehension including the story exams (N=12) and listening (N=9). As for the test task types in the reviewed tests recorded for this study, there were 10 different test task types used and the most common one was multiple-choice followed by fill-in-the-blanks and true/false as shown in Table 4.

**Table 4. Test Task Types Used in the Reviewed Tests in the Data**

| Test Task type                       | No |
|--------------------------------------|----|
| Multiple-choice cloze test           | 12 |
| Multiple-choice                      | 12 |
| Fill-in-the-blanks                   | 11 |
| True/False                           | 9  |
| Short-answer                         | 7  |
| Matching                             | 6  |
| Rewrite (sentence completion)        | 5  |
| Multiple-choice rewrite and dialogue | 4  |
| True/False/Not given                 | 4  |
| Binary-choice                        | 3  |

Considering the test task types according to the sections in the reviewed tests in the data, as Table 5 shows, the use of English section included multiple-choice cloze test as the most

frequent test task type followed by rewrite items where the students completed the second sentence so that it meant the same as the first one, multiple-choice rewrite and dialogue, fill-in-the-blanks, and binary-choice. In the institution where the data for this study was collected, the multiple-choice cloze test was prepared by first (1) searching for a relevant text from authentic materials or textbooks, (2) selecting a text that encompasses most of the target grammar points or vocabulary items studied in class and determined to be tested or that provides a clear context where those points can easily be embedded and then (3) leaving those target words, phrases or grammatical structures in the text blank by taking into consideration certain testing principles (e.g., not testing two points in one sentence). The selected text is modified where necessary to make the context clearer and to make it as representative as possible of how the testing point was studied in class. Once the testing point is determined in the text, four options are provided for it so that the students can choose the option that best fits in the blank in the text.

**Table 5. Test Task Types According to the Sections in the Reviewed Tests in the Data**

| Use of English                       | No | Reading                                 | No | Story Exam   | No | Listening          | No |
|--------------------------------------|----|-----------------------------------------|----|--------------|----|--------------------|----|
| Multiple-choice cloze test           | 12 | Multiple-choice                         | 8  | Matching     | 4  | True/False         | 5  |
| Rewrite (Sentence completion)        | 5  | True/False/Not given                    | 4  | True/False   | 4  | Fill-in-the-blanks | 4  |
| Multiple-choice rewrite and dialogue | 4  | Fill-in-the-blanks (Summary completion) | 3  | Short-answer | 4  | Multiple-choice    | 4  |
| Fill-in-the-blanks                   | 4  | Short-answer (Referencing)              | 3  |              |    |                    |    |
| Binary-choice                        | 3  | Matching                                | 2  |              |    |                    |    |

As Table 5 further shows, in the data, multiple-choice test task type was also the most common one reviewed in the reading section, which also included true/false/not given, fill-in-the-blanks (summary completion), short-answer (referencing), and matching tasks. In addition, the story exams reviewed in the data consisted of matching, true/false, and short-answer test task types with the same frequency. With regard to the listening section, true/false followed by fill-in-the-blanks and multiple-choice was the most frequent test task type reviewed in the data.

It should also be noted that detailed information regarding the proficiency level of the reviewed tests in the data, and the sections and test task types used in those tests are provided for a clear description of the language test IR interactions recorded for this study and thus, to be able to give a general picture about them. Following the principles of CA, the effect of those factors on the findings were considered and discussed only when they emerged in the data and when there was an orientation to them by the participants during their talk-in-interaction in the IR sessions.

### **3.5. Data Transcription and Analysis Procedures**

The principles of CA were followed for the data analysis of this study (See Section 3.1). After collecting the data, first of all, the videos were watched repeatedly for several times to get an overview of the data and to be familiarized with it and notes were made of the main interactional patterns and features observed in each IR interaction. During this process, two sample extracts were also selected, transcribed in detail and presented at Conversation Analysis data sessions at Hacettepe University Micro-analysis Network (HUMAN) Research Centre, which is dedicated to research on social interaction in institutional settings using a conversation analytic framework, and feedback was obtained from the participants in those data sessions. A group of CA researchers working on independent projects or studies meet in those data sessions and every week one of them brings in a selected segment of his/her data on the specified date in the pre-arranged program. Then the data session begins by distributing a copy of the transcript prepared in line with the transcription conventions to each member, providing background information about the data at hand, and playing the relevant part of the video-recording several times. During this time, the participants carefully watch the video, read through the transcript while listening to the video, analyze the extract following the principles of CA, and make notes on the transcript. They are then given some time to go through their notes and finalize their points of analysis. Afterwards, each member is invited to share his/her observations on the data, offer some suggestions for the researcher on the transcript as well as the analysis procedure in terms of, for example, the patterns or an interesting phenomenon to focus on, and can raise some points for the group to jointly discuss. Additionally, the data sessions are audio-recorded for the researcher to use later as a resource for extensive analysis.

Then the whole data was transcribed without the details for unmotivated examination. In other words, the orthographic transcription of the data was carried out. During this transcription process, too, the video/audio data and the transcripts were approached from the perspective of “unmotivated looking” to be able to notice the practices, patterns, and procedures that emerged in the data. While carrying out the unmotivated examination of the data as part of the initial analysis following the transcription, the beginning and end of the interactions for the reviewing of each test item in the whole transcribed data were determined and put in the form of extracts. Those extracts began when the participants started to discuss a test item and ended when they confirmed or revised and finalized the item so that they could proceed with the next item in the test. Then those extracts were reviewed in more detail and notes were made regarding the interactional organization and features emerging in the talk-in-interaction in each extract.

Based on the observations made while watching the videos first and then transcribing and carrying out the initial analysis of the transcribed data through “unmotivated looking” and on the feedback obtained from the two data sessions, it was seen that the problematized items constituted an important part of the IR interactions. As summarized in Figure 1, in total, there were 574 items which were interactionally reviewed in groups by the teachers and the tester together and 271 of those items were problematized, constituting 47% of the total number of items in the whole data. It was observed that the IR interactions mainly focused on the problematized items and the interactions revolved around explaining and reaching a consensus on the problem as well as revising and improving those items through a collaborative process of suggestion and decision-making. It was also the reviewing of the problematized items that showed what testing principles the teachers oriented to in and through interaction and what made an item appropriate according to them as well as the kind of processes they went through in revising a language test, going beyond key check.

On the other hand, in the case of ‘unproblematized’ test items, the IR interactions took much less time than the problematized items as the reviewers simply checked their answers and provided confirmation and acknowledgement for the intended interpretation and use of the test item (i.e., what the item tests from the syllabus). In doing this, the participants sometimes gave an account for the correct answer, offered positive assessment and sought confirmation for the validity of the item. For instance, in extract 3.1., the test item that is reviewed by the teachers CI and YE is an ‘unproblematized’ item as it is confirmed by the

reviewers in the first place. The extract starts with the reviewers announcing their answers for the second item (i.e., *to understand*) in a binary-choice test of a Use of English quiz at B1 level (lines 1 and 3), which is then acknowledged by the tester SA in line 4. The reviewers' approval of this test item becomes clear in the rest of the interaction as CI offers a positive assessment in line 5 (*bu da net/this is also clear*) and YE provides an account for the item validity in lines 7 and 10 (*o zaten net geçiyodu kitapta/after all it has appeared clearly in the textbook*), which further acknowledges CI's positive assessment and at the same time displays her stance of acceptance of the item. Meanwhile, in overlap with the reviewers' turns, SA refers to the point that the item intends to test (i.e., infinitives) in the form of account-giving in line 6, ensuring shared understanding about the interpretation and use of the item. As a result, the item receives confirmation from the reviewers and the IR interaction in extract 3.1. ends quickly without further discussion and negotiation on the item.

### Extract 3.1. Hope

- 1 CI: ikide [to under[sta:nd  
*in two it is "to understand"*  
+ SA shifts gaze to YE's paper  
+ CI and YE look at their exam papers by pointing their pencils at the second item  
+ shifts gaze to YE and then back to her paper
- 2 SA: [ ° ikide °  
*in two*
- 3 YE: [iki [to unders[tand hh  
*two is*
- 4 SA: [hnm  
+ turns gaze back to her paper
- 5 CI: [bu da net  
*this is also clear*
- 6 SA: [hope (.) [da [huh infinitive olduğu için  
*because it is infinitive with "hope"*  
+ shifts gaze to CI and then back to her paper
- 7 YE: [ ° hope [to ° evet o zaten  
*yes after all it has*  
+ underlines "hope" on her exam paper
- 8 SA: [ (inaudible)
- 9 bişey [yani  
*well it is something that*
- 10 YE: [net geçiyodu kitapta.  
*appeared clearly in the textbook*  
+ SA nods head

On the other hand, in extract 3.2. (extracts 2a, 2b, and 2c, which are different segments of one particular IR interaction, are reproduced here as a whole from chapter 4), the test item that is reviewed by the teachers CI and YE is a problematized item and thus, the interaction unfolds in a way that is quite different from extract 3.1. This item is from a tense passage presented in the form of a fill-in-the-blanks task in a Use of English quiz at B1 level where the verb “*cover*” given in a parenthesis is to be transformed into its past tense form within the context provided in a text. The IR interaction starts with the identification of a problem in the item through the reviewer YE’s initiation of a preliminary inquiry in lines 1 and 2 in search of evidence for the projected problem and the other reviewer CI’s go-ahead response (Schegloff, 2007) for problematizing the item in line 5. As announced explicitly by YE in line 6, it becomes evident that the verb “*cover*” was not studied in class and thus, the students might not be able to transform it into its past tense form. At this point, an orientation is shown to the violation of a testing principle (i.e., the item should represent a specific content area covered in the course).

Once the problem in the item is stated, a change process starts in the next phase of the IR interaction to revise the item so that the problem can be eliminated (lines 11-27). This stage is characterized by the participants’ suggestion and joint decision-making activity. YE initiates the suggestion-making sequence by offering the verb “*change*” as an alternative to replace the trouble source “*cover*” in line 11. This is then followed by YE and CI’s collaborative account-giving in lines 15 and 16 in support of the new verb and the tester SA’s acceptance of the suggested change in lines 17 and 20. Just as the tester starts to make the change on her exam paper by writing the new verb “*change*” in place of “*cover*”, which further shows her uptake of the suggestion and the group’s decision, she initiates a non-minimal post-expansion sequence in line 20 where she offers a candidate understanding (Antaki, 2012; Heritage, 1984b) of the problem and elicits further elaboration on the matter at hand between lines 22 and 27. This ensures shared understanding in the group with regard to the identified problem, the addressed testing principle and the suggested change.

The IR interaction, however, does not end here when a decision is made on a change that addresses the problem because the group wants to make sure that the new change does not cause additional problems, posing a threat to the validity of the item. As a result, in the rest of the interaction between lines 28 and 42, a new phase is launched where this time the change made in the item is reviewed. The tester SA initiates the change reviewing process in

line 28 by indirectly challenging the change through a confirmation request and a display of uncertainty about the grammatical accuracy of the target sentence as a result of the replacement of the problematic verb “cover” with “change”. This leads to the recognition of a problem in the revised target sentence in the item as the group realizes that the verb “change” does not collocate with the preposition “with” in the sentence, orienting to the testing principle about the grammaticality of the item. Consequently, a new suggestion and decision-making sequence follows where the reviewer CI offers to replace “with” in the original sentence by “using” in line 32 so that the new verb can fit grammatically well in the target sentence. She then provides an account for her suggestion in line 35, which is met with the group’s agreement and uptake of the change as SA notes the change on her copy of the exam paper in line 39 while offering a (re)formulation (Heritage & Watson, 1979) of the suggested change to demonstrate her understanding. The revision of the item is eventually finalized and the reviewing interaction for this particular item ends.

### Extract 3.2. Cover

- 1 YE: mesela cover'ı (.)  
*for example "cover"*  
 + looks at the item with her finger pointed at it and SA looks at YE's exam paper
- 2 hiç er CI hatırlıyor musun derslerde böyle  
*CI do you remember if ever in classes like*  
 + directs gaze towards CI and CI directs her gaze towards the item on her paper
- 3 hani cover'ın (.) past tense'i geçti mi?  
*you know the past tense of "cover" was encountered?*  
 + SA directs gaze towards CI
- 4 (0.3)
- 5 CI: yok ((shakes her head)) [öyle bişey geçmedi ↑de ]  
*no such a thing did not appear*  
 + directs gaze towards YE
- 6 YE: [çok onu bulamayabilirler SA]  
*they might not be able to find it SA*  
 + directs gaze at SA and then her own paper
- 7 hani er dı dı  
*well*  
 + produces thinking face and puts her hand on her forehead  
 + SA directs gaze towards YE and CI looks at her paper
- 8 (0.3)
- 9 SA: [şey mi hani-]  
*is it well*  
 + CI directs gaze towards YE

- 10 YE: [daha iyi bi ] alternatif olabilir  
*there can be a better alternative*
- 11 ama> change falan olursa  
*but if something like "change" is possible*  
+ directs gaze towards CI  
+ SA turns gaze to her paper
- 12 [belki] [change]=  
*maybe*
- 13 CI: [hu:h ]
- 14 SA: [hu:h ]
- 15 YE: =geçmiştir [diye düşündüm emin değilim ama ]  
*appeared I thought but I'm not sure*
- 16 CI: [tabi daha iyi bilirler change'i]  
*of course they would know "change" better*  
+ turns gaze to her exam paper by nodding head  
+ nods head
- 17 SA: change [diyelim mi? o zaman ona]=  
*then shall we make it "change"*  
+ YE shifts gaze to SA's exam paper
- 18 CI: [evet change'i daha iyi ] [bilirler]  
*yes they would know "change" better*
- 19 YE: =[evet ]  
*yes*
- 20 SA: <sup>1#</sup>tamam (0.2) <yani hani> verb ikisini bilemezler (.)  
+ CI writes the new verb  
"change" on her paper
- 21 [mi acaba diye<sup>#1</sup> diyosunuz]  
*okay well you know do you say so because they may not know its verb two*  
<sup>#1</sup> (lines 11-12) SA makes the change on her exam paper by writing down the new verb "change"
- 22 YE: [evet aynen öyle ] [hani-  
*yes that's right well-*
- 23 CI: [(<sup>o</sup>bunu: er<sup>o</sup>)  
*this*
- 24 SA: [tamam hocam  
*okay miss*
- 25 YE: benim ↑derdim o [change]  
*that's my concern, that "change" is*  
+ SA shifts gaze to CI and then to YE
- 26 SA: [tamam ]  
*okay*
- 27 YE: olmasa bile belki başka-  
*even if not possible, maybe something else*  
+ turns gaze back to her paper



- 28 SA: [change san]ki (.) >anlam olarak da< [oturur]  
*"change" seems to fit in terms of meaning too*  
 + directs gaze to CI  
 + YE shifts gaze to SA
- 29 CI: [°ş:ey° ] [er: ]  
*well*  
 + looks at her paper
- 30 SA: gibi di mi? [=sıkıntı (.) olmaz ]  
*like doesn't it? it wouldn't be a problem*
- 31 CI: [şey diyebiliriz ]  
*we can say this*  
 + YE turns gaze to CI
- 32 belki changed the inside of his taxi using it  
*maybe*  
 + reads from the target sentence with a pen in her hand  
 + SA shifts gaze to her paper  
 + YE turns gaze to her paper  
 + SA shifts gaze to CI and then back to her paper
- 33 falan gibi  
*something like that*  
 + shifts gaze to SA and YE
- 34 (0.3)
- 35 hani [changed with] it bi garip oluyo.  
*well "changed with it" sounds weird*
- 36 SA: [taxi ]  
 + reads from the target sentence on her paper  
 + positions hand on her paper in an attempt to note the suggested change
- 37 (0.2)
- 38 YE: huh tamam  
*okay*  
 + SA crosses out "with" on her paper  
 + shifts gaze to CI
- 39 SA: ta↑m[am with] yerine o zaman [using] diyelim  
*okay we shall say "using" instead of "with"*  
 + writes "using" in place of "with" on her paper  
 + YE turns gaze back to her paper
- 40 YE: [evet ] [evet ]  
*yes yes*  
 + CI nods head, turning gaze back to her paper
- 41 o (.) [cümleye uya- uyarlamak için e[vet =yani]  
*to adapt it for that sentence yes I mean*
- 42 SA: [tamam ] [tamam ]  
*okay okay*

The micro-analysis of this IR interaction has shown that the reviewing of problematized items involves a much more complicated process than ‘unproblematized’ items although it is seen to follow a particular sequential order. As extract 3.2. has illustrated, the reviewing of the problematized items requires a critical eye, problem-solving skills and the ability to employ various interactional actions such as explaining the problem based on evidence, offering suggestions for change in the item, reviewing the decided change critically, orienting to testing principles, displaying a situated understanding of how a language test item should be and jointly making decisions in and through talk-in-interaction. For instance, during the unmotivated examination of the data, it was observed that in those IR interactions, the participants identified, discussed and tried to overcome the following kinds of problems in the language test items;

- the item has more than one possible correct answer,
- the word, grammatical structure or expression used in the item or in the text is not studied in class and thus, students may not know and understand it,
- the item or the part of the text that is relevant for the item is too difficult or easy for the students’ level,
- the item or the text includes an ambiguous and a misleading expression or sentence that might be open to different interpretations, causing the item to be answered incorrectly,
- the word, grammatical structure, or expression used in the item or in the text is not well-formed and is inappropriate within the given context,
- the item tests a detail in the story and students may have difficulty remembering the details from the story,
- the point was already tested before or is tested twice in the same exam,
- the item has a weak and an irrelevant distractor, which is too easy to eliminate,
- the information presented in the correct option does not correspond well with the information given in the text,
- the expected answer is not the correct answer,
- there is not enough contextual clue for the correct answer,
- the question root/stem is not well-structured,
- the question is too broad and general, requiring students to write many things,
- the options are not similar in length and structure,
- the item has a clue to another item in the test or there is a clue to the correct answer somewhere in the text,

- two points are tested in one single sentence in the multiple-choice cloze test,
- the item can be answered with general knowledge, without a need to listen or read,
- the item has more than one focus,
- the expression in the item is the same as the one used in the text and is not paraphrased,
- the same structure is repeated in the options of a multiple-choice rewrite item,
- the point tested in the item is not in line with the aim of the test task,
- the prompt in the rewrite item (sentence completion) is not appropriate.

The identification of such problems in collaboratively reviewing a language test is vital for language test construction before the test is finalized. Therefore, IR interactions in the case of problematized items require greater professional skills and a systematic approach to be able to successfully deal with the problem in the item and establish mutual understanding and agreement in the IR group. Thus, it was thought that it would be worthwhile to examine the micro-analytic details of how the problems in the test items emerged in and through talk and agreed by the group, how the participants interactionally discussed such problems in the test items, what testing principles they oriented to in doing this and what procedure they followed in improving the problematic test items. By examining the interactions for the reviewing of problematized items in particular, this study can offer a comprehensive framework for language test IR interactions so that the practitioners can obtain the maximum benefit from these sessions. As a result, based on the findings of the unmotivated examination of the data, a collection was made of all the problematized items in the data (i.e., 271 items) and each extract was given a name.

The close analysis of the reviewing interactions on the problematized test items in the collection further showed that they were made up of particular phases as illustrated in extract 3.2. Thus, in the case of problematized items, three recurring phases were identified, which were also confirmed at the Conversation Analysis data sessions. They were (1) problematizing the item (PI), (2) suggesting a change (SC) to revise the item, and (3) reviewing the change (RC) made in the item (See Figure 1). These emerging phases made up the overall interactional organization of the reviewing of problematized items. As a result, a decision was made to carry out a micro-analytic investigation of the ways the three emerging phases (i.e., PI, SC, RC) were interactionally realized in language test IR sessions. Based on the findings of the initial analysis of the problematized items, sub-collections that included

the three phases of PI, SC and RC were created and the research questions, which were data-driven, were formulated accordingly. In developing the three sub-collections, the start and end of each phase in the extracts were first determined. The initial phase of PI began when the participants started the IR interaction by initiating a discussion on the problem they had identified in the item and ended when the group reached a consensus and a decision about the raised issue. The second phase of SC began with the participants' initiation of a suggestion-making sequence to revise the problematized item when they agreed with the problem and ended when the group jointly decided on a change and the tester noted it on her exam paper. The final phase of RC began when the participants this time oriented to the revised item in an attempt to recheck its appropriacy after it had undergone a change and ended when they fully confirmed the item.

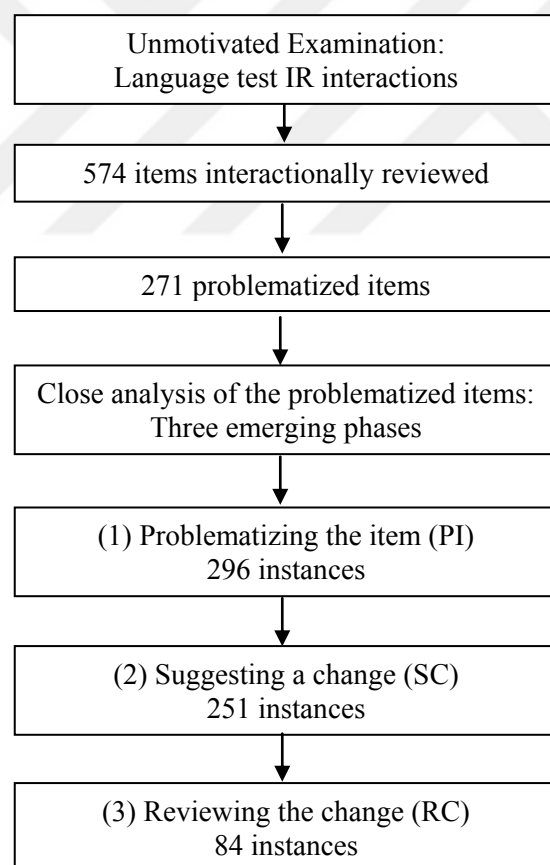


**Figure 1. Results of the Unmotivated Examination and Initial Analysis of the Data**

As Figure 1 shows, the phase of PI constituted the cornerstone of the talk-in-interaction in the reviewing of problematized items with 296 instances. There were 251 instances of SC, which stemmed from PI, and 84 instances of RC, which stemmed from SC. This means that 85% of the instances of PI were followed by a suggestion-making sequence to revise and improve the problematized items in phase SC and the changes agreed and made in 33% of the instances of SC were reviewed and further edited where necessary in phase RC. The number of instances of SC and RC was less than those of PI since after being problematized, some of the items were decided to be omitted or kept unchanged when the IR group challenged and disagreed with the presented problem. In the other few cases, the item was just problematized as the revision of the problematized item was left to the tester especially when there was not enough time to change the item and thus, SC and RC could not be employed in the reviewing interactions of those items. There were also fewer instances of RC than those of SC since the changes made in an item were not always subject to further reviewing by the group.

It should also be noted that the number of PI instances exceeds the total number of the problematized items in the collection because some items had more than one problem going through the problematization process more than once. Similarly, SC took place more than once in the case of some items as one change led to another change and different parts of an item had to be changed separately to overcome the problem (e.g., SC for the stem first and then for the options) and as items with more than one problem were subject to SC more than once. Likewise, in some cases, more than one RC was found to be employed in one single IR interaction when one RC was followed by another RC in order to review the second change made in the item or to review the same change again before finalizing the item and this process could continue until the change was fully accepted and the revised item was approved by the whole group.

Following the initial analysis of problematized items, representative extracts were selected from the three sub-collections of PI, SC and RC. Except for the extracts where the IR interaction did not continue into the other phases and ended in phase PI when the group decided to omit the item or disagreed with the problem (i.e., extracts 11a and 12a), the SC and RC of the already selected extracts from the sub-collection of PI were analyzed and presented as different segments of the same IR interaction by labelling them with alphabetical letters (e.g., 1a, 1b, 1c) so as to demonstrate the ongoing process and eventually

the complete picture of an IR interaction. In this way, the analysis could show the progressivity in an IR interaction by revealing how a particular item that is problematized in phase PI undergoes change in phase SC and how this change is then reviewed in phase RC. Thus, the representativeness of the extracts in the sub-collections of SC and RC was also considered while initially selecting extracts from the sub-collection of PI. The selected extracts were then transcribed in detail using Jefferson's transcription notation system (See Appendix B). According to Hepburn and Bolden (2013), Jefferson's transcription system "provides a shared, standard system for rendering talk-in-interaction in a way that can be textually reproduced" and "it is compact, transportable and reproducible", providing easy access to the data (p.75). Four of the selected extracts were also presented at Conversation Analysis data sessions to support their analysis.

As for the transcription procedure, besides watching the video-recordings of IR interactions to be able to capture all the details displayed in the participation framework (e.g., multi-modal resources), the Audacity Software Programme, which presents the audio track in a waveform view, was used for the acoustic analysis of the data in a more controlled way. It was especially helpful in identifying the segments of overlapping talk and marking the onset and offset of overlaps, listening to a selected segment of talk repeatedly by highlighting it and activating the "loop play" feature as well as determining the duration of pauses and silences. Furthermore, multi-modal actions that emerged as an important part of the IR interactions were marked in the transcription of the data. Due to the task-oriented nature of IR interactions where the focus is on a written product (i.e., language tests), this especially included the transcription of the participants' gaze behaviour such as looking at the test material, shifting gaze to the other participant and to each others' papers, changing gaze between the different parts of the exam paper, and maintaining mutual gaze. Other non-verbal actions emerging in relation to the written product evaluated in the talk-in-interaction were also specified in the transcription of the data. For instance, they included actions such as pointing at the item or at a key word/sentence with a pen, doing silent reading, scanning through a text, underlining or circling a key word, moving pencil or finger along a sentence/paragraph, turning over the page, crossing out the problematized part of the item, noting a change on the exam paper, writing down on paper, searching through the textbook or the story book and producing thinking face.

The multi-modal behaviours were often marked using a plus (+) sign as indicated in Jefferson's conventions. However, this plus sign only allows the marking of the onset of a non-verbal action and does not show where it ends. The need to mark the continuity of a multi-modal behaviour emerged in cases when, for example, the tester noted a decided change on her exam paper accompanied by other actions and when the participants sustained their gazes at the test material during an interval in the ongoing talk-in-interaction. Therefore, to be able to show the continuity of such non-verbal actions and to represent them accurately in the transcripts, a convention developed by Balaman and Sert (2017) was adopted to mark not only the onset but also the offset of a multi-modal action where needed. It involves the use of a hashtag sign at the onset (number + #) and offset (# + number) of a non-verbal action by numbering each set of hashtag sign used for a particular action subsequently in the same extract. Finally, in the transcripts, the participants' names were replaced with pseudonyms and letters were used to represent them so that their identities could not be recognized.

The transcription process was then followed by the CA analysis of the selected extracts to micro-analytically explain the sequential patterns and interactional actions emerging in each phase of the reviewing interactions on problematized language test items and thus, to be able to show how a problematized test item was treated in and through interaction. Since the primary aim of this study is to uncover the interactional organizations, practices and procedures that shape the talk-in-interaction in reviewing the problematized test items, the analysis involved a micro-analytical focus on turn-taking organization, sequence organization, repair and turn design as the basic mechanisms of talk and the primary units of analysis as well as the dimensions of overall structural organization, lexical choice and forms of interactional asymmetry, which are considered to be relevant places for investigating the institutionality of interaction (Drew & Heritage, 1992) (See Sections 2.6 and 2.6.1). Finally, the findings were discussed in terms of their implications for language test construction, EFL teachers' professional competence, and L2 teacher education and in terms of their contributions to institutional talk (See Chapter 5).

The procedure followed in this study for data analysis can be summarised as follows in Figure 2:

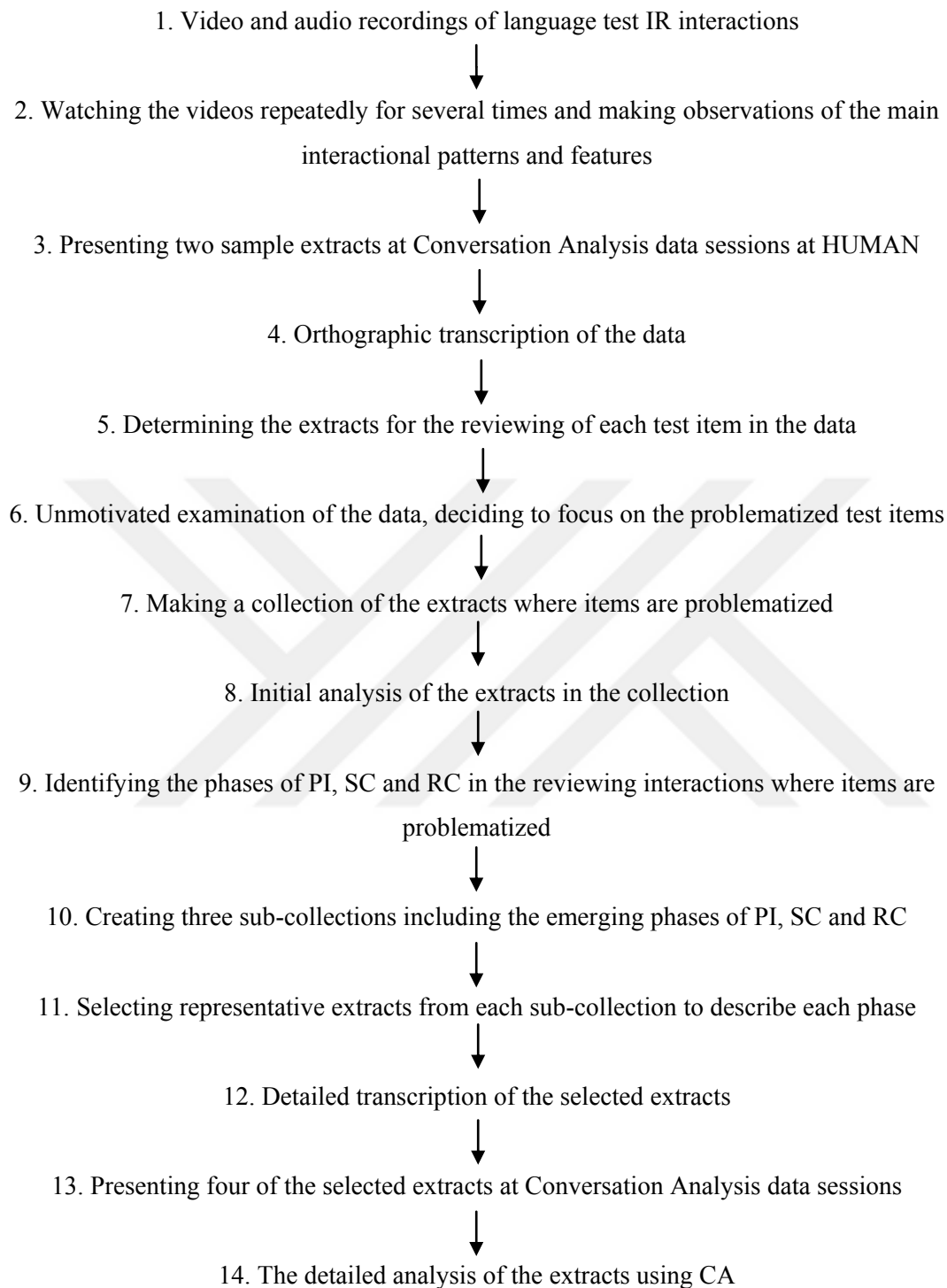


**Figure 2. Procedure Followed in the Data Analysis**



## CHAPTER 4

### DATA ANALYSIS AND FINDINGS

#### 4.0. Introduction

This chapter presents the micro-analysis of the phases of PI, SC and RC employed in the reviewing of problematized language test items in and through interaction. Each section focuses on one of the phases, beginning with PI and continuing with SC and RC respectively as this is the order of the phases that appear in the IR interactions. In this way, each phase can be scrutinized in detail while at the same time demonstrating the interactional procedure followed in reaching a mutually acceptable final resolution in dealing with the problematized items and thus, presenting a vivid picture of those language test IR interactions. Each section ends with a summary of the findings.

#### 4.1. Phase 1. Problematizing the Item (PI)

In the first extract, the two teachers BA and DE, who teach at B2 level, review a multiple-choice reading comprehension item in a level achievement test of B2 groups and provide their feedback to the tester EK. The reading comprehension item has a negative root (i.e., *Which of the following is NOT one of the advantages of the Kalenjini?*) and has four options. In extract 1a, which is the first segment of the IR interaction, BA problematizes option “b”, which is the correct answer, as she believes that it is too obvious and easy to spot for their group of students because it is not at all related to the content of the reading text and can easily be regarded *not* as an advantage of the Kalenjini. The other participants DE and EK agree with the problem presented with regard to the correct option “b”. This extract was also presented at a data session at HUMAN Research Centre. The analysis provided here is, thus, supported by the analysis carried out at this data session.

### Extract 1a. Special Diet, Segment a

- 1 DE: iki (.) "b:"  
**two is "b"**
- 2 BA: iki "b:" (.) [fakat ]  
**two is "b" but**  
+ all look at their papers
- 3 DE: 1#[di mi?]  
**isn't it?**
- 4 BA: er:: ↑ben (.) bu iki (.) "b"de (0.7) bu (.) designed  
5 by experts (.) olduğu için "↓b" (.) zaten.  
**I, for this two "b", after all it is "b" because of  
this "designed by experts"**
- 6 EK: huh huh [ °öyle bişey yok° ]  
**there is no such thing**  
+ shifts gaze to BA
- 7 BA: [er::: ] böyle bişey yok. (0.2)  
**there is no such thing**
- 8 yani (.) >bunun yerine başka bişey düşünebilir<miyiz  
9 dedim.  
**I mean I thought whether we could think of something  
else for this**
- 10 (0.2)
- 11 çünkü ↑zaten (0.2) biz#1  
**because after all as we**  
#1 (lines 3-11) DE and BA sustain gaze at their papers
- 12 bir er tribe gibi bişiyden bahsederken  
**talk about something like a tribe**  
+ shifts gaze to EK then DE  
+ DE shifts gaze between the reading text and the item on her paper
- 13 (0.4)
- 14 birden bire zaten [expert'ten söz etmek]  
**after all talking about "expert" all of a sudden**  
+ shifts gaze back to EK  
+ DE and EK sustain gaze at their papers
- 15 EK: [olmuyacağı çok açık ]  
**it is too obvious that it is not  
the one**
- 16 (0.4)
- 17 DE: [ °evet° ]  
**yes**  
+ DE and BA gaze at each other
- 18 BA: [çok ] (.) er: çok (.) böyle (0.4) açık olmuştur.  
**it has become too, too like obvious**
- 19 (0.3)

20 [hani onun yerine]  
*well instead of that*  
 + all look at their papers  
 21 EK: [ ° redundant m1 ° ]  
*is it redundant?*

The reviewing of the multiple-choice reading comprehension item starts with DE giving the correct option of “b”, which BA agrees with in line 2 by repeating DE’s words and displaying alignment while all the participants are looking at their exam papers. However, upon BA’s micro-pause following a prolonged response (“b:”), DE asks for confirmation in line 3 and in an overlap with this, BA continues her turn with “fakat/*but*” which signals an upcoming opposition to option “b” and a negative evaluation of the option. However, before BA announces the problem in the option, she interpolates a parenthetical sequence into her as-yet-incomplete turn in lines 4 and 5 with a disengaged gaze. In this parenthetical insert, which is preceded by an elongated hesitation marker, she explains with frequent pauses why the answer is supposed to be option “b” by saying that it is because of the expression “designed by experts”, which she displays as assumed knowledge as evident in her use of the word “zaten/*after all*”. This is met with a responsive uptake by the tester EK, who shows agreement with the acknowledgement token “huh huh” and then provides an account for the correct option that is complementary to BA’s explanation (°öyle bişey yok°/*there is no such thing*) by shifting her gaze to BA. She makes it clear that the information in option “b” is not present in the text and thus, it is not one of the advantages mentioned in the text, making the option correct. In line 7, BA displays alignment and agreement with EK and confirms her through a full modified repeat of EK’s first claim (Kim, 2002; Stivers, 2005).

In line 8, as they all sustain their gaze at their papers, BA requests a change for the correct option, which further projects a problem in the option, and after a brief gap of silence, she eventually problematizes option “b”, providing a justification for her request for change. While doing so, BA shifts her gaze to EK and then to DE in line 12, and turns back to EK in line 14, orienting to the other co-participants. Just before BA completes her turn of problematizing, EK overlaps with BA’s turn in line 15 and uses anticipatory completion (Lerner, 1996a) at the point where BA has finished providing a preface and the necessary background information ready to explicitly state the problem or the violated testing

principle. EK, thus, completes BA's turn with the actual problem (*olmuyacağı çok açık/it is too obvious that it is not the one*) that BA has been trying to articulate and pre-empts the projected problem by anticipating it. In this way, she shows a display of understanding, which DE also agrees with in line 17 by engaging in mutual gaze with BA. It now becomes clear that BA wants the option to be more challenging as it is too obvious and too easy to identify. According to her, since the text is about a tribe, talking about experts in the answer makes it too irrelevant and easy to answer. Then in line 18, BA continues her turn in alignment with EK's anticipatory completion through an upgraded second assessment of the option as she repeats EK's candidate evaluation of the option with inserted intensifiers (*çok/very*) and with an emphasis on the final syllable of the modified verb "*olmuş/has become*", which shows a full agreement with EK (Pomerantz, 1984a). However, the frequent pauses and the elongated hesitation marker "*er:*" as well as the use of the discourse marker "*böyle/like*" in her formulation of the negative assessment from second position at the same time signals BA's search for an alternative wording or explanation. In line 21, EK also silently asks a question as another possible explanation for BA's disapproval of option "b", offering a candidate understanding (Antaki, 2012; Heritage, 1984b) of what BA meant with regard to the problem and requesting confirmation. However, BA does not orient to EK's question and the segment ends with all the interlocutors agreeing with BA's problematization of the option and request for a change and with their orientation to the test material.

The first segment of the extract for the reviewing of a multiple-choice reading comprehension item shows how a language test item is problematized, which is the first step in the test IR interactions in the case of problematized items. The extract displays that the problem in the item is not just directly stated but rather some steps are followed to clearly explain and provide evidence for the problem so that the group can fully agree on it. For instance, before straightforwardly sharing the problem she has identified, the reviewer BA first uses a parenthetical insert where she explains what makes option "b" the correct answer and makes sure that they chose this option for the same reason, so she ensures shared understanding and common ground among the group. In this way, she also sets the ground for problematizing the option. Then she requests a change in the option, which also provides a preface that projects the announcement of a problem and gives the other participants some time to further think about the option as they all look at the item and the reading text on their

exam papers and are fully occupied with the test material. Following her request for change, BA starts to provide a justification for it by giving evidence from the reading text, comparing the topic of the text with the issue mentioned in the option. At this point, BA's shift of gaze from her exam paper to the co-participants also draw their attention to the item as DE starts to shift her gaze between the reading text and the target item, and EK becomes more focused on the test material by sustaining her gaze at it. Those steps lead to the collaborative identification of the problem and its realization and acceptance as EK eventually uses anticipatory completion (Lerner, 1996a) and states the actual problem pre-emptively, orienting to the violated testing principle that the answer is too obvious and offering a candidate evaluation of the option before BA does so. This also illustrates that not only the problematizer but also the other participants are active agents in the IR interaction; in other words, they all actively and critically think about the item as prompted by BA's talk and attempt to make a contribution to the process of problematizing the item. Although DE is relatively more passive than the others, her orientation to the test material, her sustained gaze at her paper throughout the segment and her shift of gaze between the reading text and the item on her paper in line 12 show her engagement in the task, her constant thinking about the matter at hand and her acceptance of the problem as she also responds with an agreement token in line 17. According to Goodwin (1981), different from ordinary conversation, engagement in a competing activity and looking at an object used for the accomplishment of an activity that competes with the talk is less sanctionable than looking away in general and is not a sign of diminished engagement in the conversation.

This extract also shows an orientation to the institutional role distribution of the participants. This is especially evident in the interaction between the tester EK and the reviewer BA. First of all, upon BA's explanation for why option "b" is the correct answer, EK's provision of an elaborative account in line 6 in the form of a declarative statement signals her right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item as the writer and thus, the 'owner' of the item. Similarly, BA's shift of gaze to EK while problematizing the item in lines 12 and 14 and addressing the problem to her in particular orients to EK's role as the tester who has prepared the test and thus, to her right to learn about the problem identified by the reviewers. Secondly, BA's full modified confirming repeats (Stivers, 2005) of EK's assertions in lines 7 and 18 display an orientation to her responsibility as the reviewer who has identified the problem in the item. In line 7, for instance, she repeats EK's full claim rather than simply confirming it through an agreement

token to show that she is well aware of the matter at hand and knows what makes option “b” the correct answer and that the problem in the option is not simply about that. Similarly, by engaging in an anticipatory completion, EK carries out the first position negative assessment of the option and thus, goes first in announcing the problem. As a result, in line 18, BA modifies EK’s prior evaluation of the option by inserting intensifiers and adding an emphasis to the final syllable of the verb in her turn-constructural unit (TCU). By employing modified repeats, BA does not only show full agreement with EK through upgraded second assessment (Pomerantz, 1984a) but also claims her epistemic right to confirm her and asserts epistemic authority and control as the reviewer and the problematizer of the option (Heritage & Raymond, 2005; Stivers, 2005). In this way, BA indexes her “own primary rights to have made the claim originally offered by the first speaker” (Stivers, 2005, p.141) because as the first assessor of the option, EK “implies a claim of primary epistemic and/or moral rights to assess that state” (Heritage & Raymond, 2005, p.34). Although both BA and EK have equal access to the state of affairs (i.e., the test material, reading text and the item), by upgrading her second position assessment through modified repeats, BA orients to her right to announce the problem and thus, to her institutional role as the reviewer of the test and the problematizer of the item.

Extract 2a demonstrates how a language test item is problematized collaboratively. In this extract, the two teachers YE and CI teaching at B1 level review an item in a grammar quiz and provide their feedback to the tester SA. They focus on an item in a tense passage presented in the form of a fill-in-the-blanks task where verbs are given in parenthesis to be transformed into their right tenses within the context given in a text. YE problematizes the verb given in one of the blanks to be transformed into its past tense form as she claims that the verb which is “*cover*” is not studied in class and students may not have encountered it in any of the exercises done in class. CI and SA agree with the presented problem. This extract was also presented at a data session. The analysis provided here is, thus, supported by the analysis carried out at this data session.

#### Extract 2a. Cover, Segment a

- 1 YE: mesela cover'ı (.)  
     *for example "cover"*  
     + looks at the item with her finger pointed at it and SA looks at YE's exam paper
- 2 hiç er CI hatırlıyor musun derslerde böyle  
     *CI do you remember if ever in classes like*

- 3 + directs gaze towards CI and CI directs her gaze towards the item on her paper  
hane cover'ın (.) past tense'i geçti mi?  
***you know the past tense of "cover" was encountered?***  
+ SA directs gaze towards CI
- 4 (0.3)
- 5 CI: yok ((shakes her head)) [öyle bişey geçmedi ↑de ]  
***no such a thing did not appear***  
+ directs gaze towards YE
- 6 YE: [çok onu bulamayabilirler SA]  
***they might not be able to find it SA***  
+ directs gaze at SA and then her own paper
- 7 hane er dı dı  
***well***  
+ produces thinking face and puts her hand on her forehead  
+ SA directs gaze towards YE and CI looks at her paper
- 8 (0.3)
- 9 SA: [şey mi hane-]  
***is it well***  
+ CI directs gaze towards YE
- 10 YE: [daha iyi bi ] alternatif olabilir  
***there can be a better alternative***

The IR starts with YE's initiation of a pre-expansion sequence for item problematization in lines 1-3. As she is looking and pointing at the target item on her exam paper, she sets a ground for problematizing it by directing a question at the other reviewer CI in the form of a *do you remember* recognition check (You, 2015) and asking her whether she remembers if the past form of the verb (i.e., "cover") given in the item was studied in class. At this point, the tester SA shifts her gaze to CI and CI directs her gaze towards YE, providing a negative response (*yok/no*) in line 5 and then explaining with reference to the past instructional events that the verb "cover" was not encountered in class (*öyle bişey geçmedi ↑de/such a thing did not appear*), which acts as a go-ahead response for problematizing the item (Schegloff, 2007). Only when YE gets a response from CI does she state her own assertion regarding the problem in the test item in an overlapping turn with CI's response turn.

In this base first pair part (FPP) of problematizing the item in line 6, YE directs her gaze towards the tester and explicitly addresses her in turn-final position as she is giving her own account of the problem in alignment with CI's response. In line 7, she retains the floor in an

attempt to provide further explanation as evident in her use of a hesitation marker and the discourse marker “*hani/well*” as well as her display of a thinking face by putting her hand on her forehead while at the same time looking at her exam paper. After a gap of brief silence and YE’s delayed production, SA takes a turn in line 9 and attempts for an anticipatory completion (Lerner, 1996a) in the form of a candidate understanding (Antaki, 2012; Heritage, 1984b), trying to utter the explanation YE has been searching for and at the same time check her understanding of the problem just mentioned. However, just as SA takes the turn, YE produces her explanation in line 10, cutting off SA’s utterance. According to her, a better alternative can be found for the verb, which is considered to be problematic. This projects the change process for the improvement of the item as the relevant next action or phase in the IR interaction.

This part of the extract demonstrates the act of problematizing an item as the first phase in a language test IR interaction. More specifically, it shows the problematizer YE’s attempts at proving the problem and reaching a consensus among the group so that the item can be improved. The problem YE identified is the inappropriacy of the verb given in one of the blanks (i.e., *cover*) to be changed into its past form since it was not studied or encountered in class. The extract demonstrates that problematizing a language test item in an interaction with the tester and the other teachers as item reviewers is not simply about stating the problem but it involves a dynamic process of providing account and evidence from various perspectives and convincing everyone in the group of the problem. Although the problematizer had enough time to think about the item and her past teaching experience before the IR interaction, she does not, for instance, directly state her assertion regarding the problem but instead gets engaged in a preliminary inquiry (Pomerantz & Heritage, 2013) by addressing a question at the other reviewer to check out her experience with respect to the teaching of the verb “*cover*” and to make sure that there is enough evidence for the problem. More specifically, YE seeks evidence by using a recognition check with *do you remember X* construction and making reference to the past instructional events (Can Daşkın & Hatipoğlu, 2019; You, 2015) in order to find out if the past tense of the verb “*cover*” was encountered in class. This evidence is acknowledged with CI’s negative response to the preliminary inquiry, which encourages the production of the projected problem in the item, acting as a go-ahead response (Schegloff, 2007) and suggests an upcoming acceptance of the problem likely (Pomerantz & Heritage, 2013). Thus, only after such a preliminary inquiry in the pre-sequence, YE does the actual act of problematizing, which seems to be welcomed by the



tester SA. In a way, through this recognition checking question, YE invites a collaborative identification of the problem and manages to involve all the co-participants in problematizing the item as in extract 1a. For instance, in line 9, even the tester SA makes an attempt to explain the problem YE has been trying to articulate, displaying alignment. Thus, in and through the talk-in-interaction, they provide a justification for the problem in the item with reference to the shared past instructional events (Can Daşkın & Hatipoğlu, 2019; You, 2015) and collaboratively show an orientation to the testing principle that was violated (i.e., that the item tests what was not taught). In both extracts 1a and 2a, the problematizers use some interactional resources (e.g., YE employs preliminary inquiry with a recognition check, BA provides evidence and the necessary background information) that lead the other participants to identify the violated testing principle and announce the problem before they do so, which eventually results in the direct acceptance of the problem in the group. This further shows that there are certain testing principles that all the co-participants in the IR group adhere to in the construction of a language test and that they share a common ground in this respect.

As in extract 1a, this extract also shows an orientation to the institutional role distribution of the participants. YE directs the question to CI, not SA because as the reviewers and teachers teaching at the same proficiency level following the same program, CI is in the same epistemic position as YE and they share the same institutional responsibilities. However, while problematizing the item in line 6, YE explicitly addresses SA this time, orienting to her role as the tester who has prepared the test and thus, to her right to learn about the problem identified by the reviewers so that they can work towards resolving the problem in the test. At this point, SA is also ascribed a K- epistemic position (Heritage, 2013) as she does not have an equal access to the domain of information that the teachers YE and CI have because they are the ones who actually teach the content of the test and who have greater command of the teaching process. In fact, this is the goal of the IR interactions with the teachers, that is to create the dynamic interplay between teaching and testing. Hence, YE turns to SA and presents the problem from the students' perspective (*bulamayabilirler/they might not be able to find*) by explaining that they would not be able to change the verb “cover” into its past form. Upon this, although she cannot achieve it when YE cuts off her utterance, in line 9, SA attempts to display her uptake of the problem stated by YE by employing an anticipatory completion in the form of a candidate understanding, which is used as an affiliating device (Hayashi, 2013).

Extract 3a demonstrates another case of problematizing a language test item in and through interaction. In this extract, the two teachers BA and DE teaching at B1 plus level review an item in a use of English quiz prepared for B1 plus groups and provide their feedback to the tester ME. They focus on a grammar item in a multiple-choice cloze test which tests students' knowledge of the conjunctions they have studied in class. They are to circle the conjunction among four alternatives (i.e., *because*, *therefore*, *furthermore*, *however*) that best fits in the blank and the context that is given in a text. The expected answer is "*furthermore*"; however, both reviewers problematize the item, claiming that "*therefore*" and "*however*" are also possible as correct answers. In the end, the group reaches a consensus on the problem. This extract was also presented at a data session. The analysis provided here is, thus, supported by the analysis carried out at this data session.

### Extract 3a. Therefore, Segment a

- 1 DE: [şimdi: beşte: ]  
*now for five*  
+ all gaze at the item on their exam papers
- 2 BA: [huh şu (.) beşe gelelim] ikimiz de aynı şeyi  
+ points at the fifth question on her paper
- 3 (0.3)
- 4 işaretlemiştir.  
*let's come to this number five, we both marked the same thing*  
+ shifts gaze to DE's paper
- 5 DE: therefo:re  
+ shifts gaze to BA's paper  
+ ME shifts gaze to DE's paper
- 6 ME: therefore'u mu işaretlediniz hocam  
*did you mark "therefore" miss*  
+ all look at their papers
- 7 BA: ((laughs))=
- 8 ME: [((laughs))]
- 9 DE: =e↑vet [bu bunun sonucu çünkü ↑further[more da]  
*yes because this is its result, "furthermore" is also*
- 10 BA: [tabi ]  
*of course*
- 11 DE: olur howe[ver da olur  
*possible, "however" is also possible*  
+ shifts gaze to ME
- 12 BA: [ikisi de oluyo?  
*both are possible*  
+ shifts gaze to ME

- 13 DE: however da?  
*"however" too?*  
 + shifts gaze to BA's paper
- 14 BA: [however ] da olur  
*"however" is also possible*  
 + shifts gaze to DE
- 15 ME: [ °öyle mi ° ] ciddi fmişin [((laughs))  
*oh really are you serious*
- 16 BA: [çünkü onların olmaması  
 için bi sebep yok=  
*because there is no reason for why they cannot be the  
 correct answer*
- 18 DE: =yok  
*no*
- 19 ((all turn their gazes to the item on their exam  
 papers and they silently read the relevant part for  
 6.8s))
- 20 DE: bak sonuç olarak bakarsan [therefore-  
*look if you consider it as a result*
- 21 BA: [↑however olmuyo=  
*"however" is not possible*
- 22 ME: =however [olmaz hocam  
*"however" is not possible miss*
- 23 DE: [huh however [değil  
*it is not "however"*
- 24 BA: [fur- ama furthermore  
*but*
- 25 ve therefore olur=  
*and are possible*  
 + shifts gaze to ME
- 26 ME: =broken [off]'dan dolayı (.) there(.)fore'a gittiniz  
*because of "broken off" you chose "therefore"*  
 + sustains gaze at her paper
- 27 DE: [fu-]
- 28 ME: di mi  
*didn't you?*
- 29 <sup>1#</sup> brok[en off [from much larger areas of [ice
- 30 BA: [tabi  
*of course*
- 31 DE: [e↑vet  
*yes*
- 32 BA: [tabi  
*of course*
- 33 ME: in the ocean because of the change in temperature<sup>#1</sup>

#1 (lines 29-33) ME reads the relevant sentence while DE and BA sustain their gaze at their papers  
 34 (0.9)  
 35 er dahası-  
**furthermore**

Extract 3a begins with the two reviewers DE and BA giving the first possible correct answer “*therefore*”. In doing this, using an inclusive language in lines 2 and 4 (*ikimiz de/we both, işaretlemişiz/we marked*) and shifting her gaze to DE’s exam paper, BA orients to the other reviewer DE and finds out that they both marked the same answer. Following this, DE reveals the answer that they have marked in elongated form (*therefo:re*) by directing her gaze towards BA’s paper. However, this is met with a question by the tester ME in line 6 in the form of a confirmation request and this, together with ME’s laughing following DE’s confirming response in line 8 displays ME’s surprise at the produced answer. This projects some kind of trouble in the item and it becomes clear that “*therefore*” was not the expected answer. In addition, BA’s laughing immediately after ME’s request for confirmation indicates BA’s understanding that there is something unexpected in the item.

In line 9, following her confirming response, which she offers with a rising intonation, DE provides an account for their choice of the answer “*therefore*” and then immediately lists the other possible correct answers “*furthermore*” and “*however*” by shifting her gaze to ME. It is at this point that DE is orienting to the violated testing principle that there is more than one possible correct answer and thus, announcing the problem in the test item. At certain points during DE’s turn-in-progress as she is listing the possible answers one by one, BA also overlaps and not only shows full agreement but also provides confirmation through the use of the marked confirmation “*tabi/of course*” (Stivers, 2011) in line 10 and a (re)formulation (Heritage & Watson, 1979) which conveys the gist of DE’s assertion with a slight rising intonation and a shift of gaze to ME in line 12 (*ikisi de oluyo?/both are possible*). This displays BA’s alignment and affiliation with the stance taken by DE in problematizing the item. In lines 13 and 14, BA and DE further draw attention to the other correct answer “*however*”, emphasizing its possibility as DE elliptically repeats her prior assertion regarding “*however*” with a slight rising intonation (*however da?/too?*) and BA confirms it by repeating it in full form with a shift of gaze to DE.

In response to the reviewers' problematization of the item, ME shows partial resistance and a display of surprise through the quietly uttered change-of-state token (Heritage, 1984b) and newsmarker “*öyle mi /oh really*” followed by the expression of surprise “*ciddi misin /are you serious*”, the final part of which is produced with a smiley voice, and by laughter. Following ME's turn of surprise and change of state, BA initiates a non-minimal post-expansion sequence in lines 16 and 17 for further elaboration of the problem (Stivers, 2013) by providing a general account that there is no reason for why the other two options cannot be possible as correct answers. Without any time lapse, DE immediately displays alignment and agreement through a partial modified repeat (Kim, 2002; Stivers, 2005) of BA's turn by using the final word of her turn (i.e., *yok/no*).

As the non-minimal post-expansion sequence continues, all the co-participants orient to the item on their exam papers and read the relevant part silently on their own in search of more concrete evidence. Then in line 20, DE attempts to further explain why “*therefore*” is possible but her utterance is cut off by BA, who in overlap, uses self-initiated self-repair with regard to the problem they stated earlier. She asserts that “*however*” is not possible with a rising intonation on the first syllable of “*however*” after finding out new evidence in the reading text against their prior claim that “*however*” is a possible correct answer. This is immediately acknowledged with a responsive uptake by ME, who repeats BA's full assertion with a greater certainty by addressing the reviewers with “*hocam/miss*” and displaying the information as assumed knowledge as evident in her use of the simple present tense (*olmaz/not possible*). This also shows ME's partial resistance to the reviewers' earlier remark regarding “*however*”. In an overlap with ME's turn, in line 23, DE also shows agreement with an acknowledgement token followed by a full modified repeat (Kim, 2002; Stivers, 2005) of BA's utterance, which reproduces her core claim. In this way, the reviewers disprove their prior claim regarding “*however*” and agree that it is not possible as a correct answer within the given context after rechecking and finding conflicting evidence for it. As a result, in lines 24 and 25, BA repairs and finalizes the problem in the item by once again shifting her gaze to ME, stating that “*furthermore*” and “*therefore*” are the possible correct answers. After eliminating “*however*”, ME then quickly orients to the other possible correct answer “*therefore*” with her gaze sustained at her exam paper and seeks further elaboration and confirmation for it by offering a candidate explanation in line 26 and then reading aloud the relevant sentence. While doing so, BA confirms ME two times with

the stance marker “*tabi/of course*” (Stivers, 2011) in lines 30 and 32, displaying certainty and DE agrees with ME’s candidate account by using a compliance token with a rising intonation in line 31. After a gap of silence, in line 35, ME thinks aloud to recheck whether “*furthermore*” also fits well in the given context but it is cut off as they move onto the next phase of SC in the IR interaction (See Extract 3b).

This extract demonstrates how the act of problematizing an item is shaped collaboratively in and through interaction and the role account-giving and evidence-seeking play in achieving this. Initially the reviewers simply listing the possible correct answers as indicative of the problem is not found sufficient and the tester ME especially makes some attempts in seeking concrete evidence so that she can be convinced of the problem. Her change of state and display of surprise in lines 6, 8 and 15, her confirmation requests in lines 6 and 26 as well as her display of partial resistance and challenging in lines 15 and 22 all contribute to the group’s careful evaluation of the options and the repair and acceptance of the problem. As a result, following ME’s turns, the reviewers provide an account for their choice of multiple answers, first for “*therefore*” (line 9) and then for all the possible correct answers in general (lines 16 and 17) and orient to the test material in search of more evidence in a non-minimal post-expansion sequence, which eventually helps them recheck their assertion regarding the selected options and repair the problem by disproving their prior claim about “*however*” and eliminating it. Thus, the talk-in-interaction and the orientation to the test material allow one of the reviewers BA to get engaged in a self-initiated self-repair in correcting the problem they stated earlier. ME’s final confirmation request in the form of a candidate explanation for the possibility of “*therefore*” as another correct answer finally ensures the group’s agreement and shared understanding on the problem. It is also seen that throughout the interaction, the accounts proceed from general (e.g., *çünkü onların olmaması için bi sebep yok/because there is no reason for why they cannot be the correct answer*) to more specific (e.g., *broken off’dan dolayı/because of “broken off”*) until the problem is finalized and a mutual agreement is reached in the IR group. This extract once again illustrates that not only the participant who has identified the problem but also all the other participants are actively involved in explaining and understanding the problem, collaboratively making a contribution to the process of problematizing the item. As in the case of other extracts, this extract also shows that problematizing an item is much more than the problematizer simply stating the trouble in the item as it requires constant checking, negotiation and compromise.

As in the other extracts, this extract further demonstrates an orientation to the institutional role distribution of the participants, giving clues about reviewer and tester roles and responsibilities. Regarding the reviewer roles, first of all, the reviewers' shift of gaze to each other's papers as they give the first possible correct answer as the sequence-initiating action in lines 4 and 5 as well as in line 13 while providing the other possible correct answer and BA's use of an inclusive language show that BA and DE as the reviewers share the responsibility for problematizing the item and want to make sure that they marked the same answers before announcing the problem and orienting to the violated testing principle. In this way, as in extract 1a where BA uses parenthetical insert to ensure shared understanding about their choice of the answer before problematizing the item, the reviewers establish common ground on which to base the problem and reveal the possible answers (i.e., the problem) only when they find out that they have approached the item in the same way. As a result, the reviewers who are in the same position orient to their experience of the test-taking process during which they had a thought on the item and to their right and responsibility to present their evaluation of the item to the tester.

With respect to the teachers' roles as reviewers, BA and DE convey similar stances while instantiating the act of problematizing. They, for instance, use modified repeats (Stivers, 2005) of each others' assertions as BA does in line 14, and DE does in lines 18 and 23. Through those repeats, the reviewers claim their epistemic right to confirm each other and thus, to assert epistemic authority and control, orienting to their role as reviewers and their right to problematize the item (Heritage & Raymond, 2005; Stivers, 2005). In addition, BA's use of the stance marker "*tabi/of course*" as marked confirmation in line 10 in response to DE's account and in lines 30 and 32 in response to ME's confirmation request shows her attempt to upgrade her epistemic stance in order to assume a superior epistemic status as the reviewer who has done her job of reviewing and evaluating the item but who could not go first in sharing her K+ epistemic perspective regarding the item (Heritage & Raymond, 2005). By using "*tabi/of course*", BA also treats the presented information as assumed knowledge, implying that the state of affairs indexed by ME's confirmation request is self-evident and BA's position on the matter is previously held (Heritage, 2002; Stivers, 2011). While BA upgrades her epistemic stance, the tester ME downgrades hers by providing an account in the form of a candidate explanation and then asking for its confirmation from the reviewers. This could be because as the first provider of specific evidence from the test material which they all have equal access to but as the one who is not

actually problematizing the item, she might not want to implicate epistemic priority and wish to hear the explanation from the reviewers' perspective so that she can understand the problem better (Heritage & Raymond, 2005).

Finally, with respect to the tester roles and responsibilities, BA and DE's shift of gaze to the tester ME while announcing the problem in lines 11 and 12 as well as in line 25 while stating the revised and finalized problem, and addressing her in particular signals the tester's right to learn about the problem identified by the reviewers so that the item she has prepared can be revised. In addition, ME's final confirmation request in the form of a candidate explanation for the possibility of "*therefore*" as another correct answer in line 26 orients to her right and responsibility to seek an account for the reviewers' choice of an option and to make sense of the problem as the tester who is accountable for the test and who is to make a decision accordingly in the rest of the interaction.

In extract 4a, the two teachers DU and GE teaching at B1 level review a multiple-choice reading comprehension question in a level achievement test of B1 groups and provide their feedback to the tester SA. The reading text that the multiple-choice question is based on is about a company named "Muzak" making specific music appropriate for different businesses. The multiple-choice question asks what happens when Muzak soundscapes are used and lists in the options the possible results of using such music. DU problematizes option "b" in this item, which states "*customers do shopping very fast*" as she believes that it could also be possible as a correct option and the item might end up having two possible answers. In the end, GE and SA agree with the problem presented for the distractor option "b".

#### Extract 4a. Muzak Soundscapes, Segment a

- 1 SA: dörtte ho[cam]  
*for four miss*  
+ all look at their exam papers
- 2 DU: [dört] ["d" di mi  
*four is "d" isn't it?*
- 3 GE: [denizli: (.)  
*it is "d" (coded with the name of a  
city in Turkey called "denizli")*
- 4 [evet more efficient kelimesinden  
*yes from the word "more efficient"*  
+ looks at the relevant part in the reading text



5 SA: [denizliyi hocam hedeflemiştim  
*I aimed for "d" miss*  
+ shifts gaze to DU's paper

6 GE: <sup>1#</sup> çıkartıyorlar  
*they derive*  
=effi[cient'ı

7  
8 SA: [evet  
*yes*

9 GE: da bi↑lirler herhalde sanki geçti di mi? efficient  
*they would probably know, I think "efficient" was encountered, wasn't it?*

10 (0.6)

11 bi bi- bundan evvelki kitaplarda geçti  
*it appeared in the previous books*

12 =gerçi calme- er er er yok<sup>#1</sup>  
*though no*  
#1 (lines 6-12) all look at their exam papers

13 (0.5)

14 personnel [are  
+ reads the relevant part in the reading text

15 DU: [ben şey de olabilir mi acaba  
+ SA directs her gaze towards DU

16 [dedim <sup>2#</sup> customers  
*I thought whether this thing would also be possible*

17 GE: [more efficient  
+ turns to the item on the next page

18 SA: [°hangisi hocam°  
*which one miss*

19 DU: do shopping very fast<sup>#2</sup>  
#2 (lines 16-19) DU reads the option with a pencil in her hand while DU and SA look at the relevant part on their papers

20 (1.0)

21 >şimdi< çünkü (.) hani (.) hızlı müzik çaldığı  
22 [yerler]de  
*now because well in places where fast music is played*

23 SA: [huh ]

24 DU: insanlar daha (.) ak- hızlı hızlı hemen  
+ DU directs her gaze at SA and moves her hands in fists forward  
and back with a repeated alternating movement  
+ SA and GE look at the reading text

25 [alışveriş] yapıyorlar ya  
*you know people do shopping faster immediately*

26 SA: [hu:h ]

27 (0.3)

28 GE: hu:h

29 DU: bu sorunun soruluş (.) tarzında da (.)  
*also in the way this question is asked*  
+ points at the stem and looks at it  
+ SA directs her gaze towards the item on DU's paper while GA scans through the text in her paper

30 when muzak >soundscapes< are used kullanıldığında  
*are used*  
+ reads the stem while SA sustains her gaze at DU's paper

31 SA: huh-[huh]

32 DU: [bi ] tanesi de o çünkü (.)  
*because that is also one of them*

33 [hızlı hızlı alış]veriş yapıp çıkıyorlar  
*they do shopping fast and then go*  
+ DU moves her hands in fists forward and back with a repeated alternating movement and directs her gaze at SA  
+ SA and GE look at the reading text on their papers

34 SA: [hu::h ]

35 GE: hu:h

36 (0.3)

37 DU: [o da doğru canım.]  
*it is also correct dear*  
+ sustains gaze at SA

38 SA: [aslında burda ] şey var uhm  
*actually here there is this thing*

39 (0.2)

40 <sup>3#</sup>in a store (.) people seem to shop in a more relaxed

41 and leisurely manner diyo ama (.) dediğiniz gibi <sup>#3</sup>  
*it says but as you said*  
#3 (lines 40-41) SA reads a sentence from the text and GE and DU follow her by looking at their papers

42 başka paragraflarda da  
*in the other paragraphs too*

43 [ona gidecek şey var]  
*there is something referring to it*  
+ shifts gaze to DU  
+ DU looks at SA and nods

44 GE: [beşinci paragrafta ] da  
*in the fifth paragraph too*

45 DU: [var ]  
*there is*

46 GE: [hızlı] [er ]  
*fast*

47 SA: [aynen] (0.4) <sup>4#</sup>hu:h  
*exactly*  
+ directs her gaze at GE

- 48 DU: çünkü genel sorulmuş ↑ya (.) when muzak (.) [loud-]  
*because you see it is asked in general terms*  
+ reads the stem
- 49 SA: [evet ]  
*yes*
- 50 DU: soundscapes are [↑u:sed kullanıldığında. ]  
*are used*
- 51 GE: [oraya da gidebilir çocuk evet]  
*yes the kid can also go for it*
- 52 SA: [doğru doğru tabi yani ]  
*right right of course you know*
- 53 öğrenci sırasıyla gitmeyip başka bi yerden çıkarcak  
54 olsa [doğru]  
*if the student does not read in an order but derive it from somewhere else, true*
- 55 GE: [evet ]<sup>#4</sup>  
*yes*  
#4 (lines 47-55) all look at their exam papers

As all the participants are looking at their exam papers, the IR starts with the tester SA directing the group's attention to the fourth item. One of the reviewers DU then initiates a pre-expansion sequence that serves the base sequence of problematizing the item (Schegloff, 2007) by soliciting confirmation for the answer that she chose. In an overlap with DU's provision of the answer, GE also gives the same answer by coding the letter of the selected option, that is "d", with the name of a city in Turkey with an elongation at word-final position (denizli:). Then in lines 4 and 5, in overlapping turns, SA and GE both offer a confirming response to DU's confirmation request. The already-speaking participant GE does it with the positive response token "evet/yes" followed by the presentation of evidence (i.e., more *efficient*) from the part of the reading text where the students are expected to derive the answer from. In doing this, she emphasizes the particular word in the reading text that has led to the answer. On the other hand, with a shift of gaze to DU's exam paper, SA offers her confirmation by pointing out that option "d" was what she aimed for in constructing the item. At the possible completion point of GE's ongoing unit of talk, SA also shows agreement via the acknowledgement token "evet/yes", displaying alignment with GE's turn of account-giving.

As all the co-participants sustain their gaze at their papers, GE continues her turn with no lapse and rushes through into another unit of talk where she plays the devil's advocate by first claiming that the students would possibly know the word "*efficient*" by downgrading her claim with the hedges "*herhalde/probably*" and "*sanki/I think*", and then seeking confirmation for it in line 9 to check the correctness of her assumption. In formulating her claim, she evaluates the word, which is key to the correct answer, from the students' (*biñlirler/they would know*) and teaching perspectives (*geçti/was encountered*) by making reference to a past instructional event (Can Daşkin & Hatipoğlu, 2019; You, 2015) where "*efficient*" is assumed to have been encountered. However, the other participants do not orient to GE's confirmation request as they are engaged in the test material, further thinking about the item and thus, a gap of silence ensues with no uptake. Then in line 11, GE takes the floor again in search of evidence for her confirmation request as signalled by her turn-initial cut-offs and then supports her prior claim by locating the evidence for the teaching of "*efficient*" in the previous books they studied in class, providing the sought-after confirmation herself. Immediately following this, GE rushes to launch a self-initiated self-repair with respect to the evidence for the correct answer by using the repair preface "*gerçi/though*" (Lerner & Kitzinger, 2010) and then orients to a different word (*calme-*) in the same sentence in the reading text that would actually present a clue for the answer. However, the word is aborted by cutting it off in the course of its production and following the use of hesitation markers, it is further repaired (*yok/no*) in the same TCU in line 12, leading to self-cancellation. After a brief gap of silence, GE turns to the sentence including "*efficient*", which she focused on previously, and reads it aloud in lines 14 and 17, once again signalling that it is the sentence where the answer lies.

As GE reads aloud the relevant sentence, having the correct answer confirmed in the pre-expansion sequence, DU overlaps and launches a turn with a preface, which creates a context for problematizing the item by signalling indirectly through an embedded question that one of the options could also be possible as an answer and that there is a problem with the item (lines 15-16). This draws others' attention as SA directs her gaze towards DU in line 15, asking quietly in an overlapping turn which one is the problematic option in line 18 and GE turns to the item on the next page in her exam paper. DU reads aloud the problematic option with a pencil in her hand while GE and SA follow her by looking at the

relevant part on their exam papers. It is at this point that DU implements the act of problematizing the item and orients to the violated testing principle that there is more than one possible correct answer. Then between lines 21-37, DU holds the floor, providing account for why she thinks the option causes trouble. In doing this, first, in lines 21-25, by referring to the context given in the reading text, she explains that in places where fast music is played, people do shopping fast. Her explanation is accompanied by a non-verbal account as she moves her hands in fists forward and back with a repeated alternating movement to demonstrate “fast” while looking at the tester, SA, who in line 23 shows listenership (huh) as she is looking at her exam paper. In turn-final position of her unit of talk in line 25, DU uses “*ya/you know*”, which displays her account as assumed and shared knowledge already expressed in the reading text, which they all have equal access to (Heritage, 2013). Just prior to the end of DU’s turn of account-giving in line 26 and at its first possible completion point in line 28, SA and GE respectively display agreement via an elongated acknowledgement token (hu:h) while scanning through the reading text on their papers.

Between lines 29-37, DU extends her turn to provide further account and evidence by this time referring to the stem of the item and pointing at it, at which point SA directs her gaze towards the item on DU’s paper and GA scans through the text on her paper. DU then reads aloud the stem in line 30 and upon reading it, she translates the verb in the final position of the root into Turkish (*kullanıldığında/are used*), which is received with SA’s display of understanding through the acknowledgement token “huh-huh” in line 31. Following this, DU makes a connection between the stem and the second possible correct option in lines 32 and 33 by explaining that the way the question is asked allows the item to be completed well with the information given in the option in question (i.e., doing shopping fast). While explaining the action of “doing shopping fast” as one of the possible outcomes of using Muzak soundscapes in line 33, DU again moves her hands in fists forward and back with a repeated alternating movement and directs her gaze at SA, who then shows agreement through an elongated acknowledgement token in line 34 by looking at her exam paper. At the transition relevance place upon DU’s completion of her account, GE also displays agreement in the same way in line 35. After a brief gap of silence, DU concludes by repeating the problem in line 37 and providing her final comment that the option is also possible as a correct answer. In doing this, DU sustains her gaze at the tester SA and addresses her with an endearment address term in turn-final position (*canım/dear*).

In overlap with DU's turn, SA this time shows disalignment in line 38 by prefacing her turn by "*aslında/actually*", which projects some kind of disagreement, and presenting counter-evidence from the text that would eliminate the option. Thus, SA exhibits a "pro forma" agreement (Schegloff, 2007) as she initially displays understanding and minimal agreement through the use of an acknowledgement token, which serves to delay her upcoming dispreferred response as a contiguity-breaking practice. But then in line 41, she quickly backs down and withdraws her counter-evidence by disproving it herself based on DU's prior account. In doing this, she displays alignment and affiliation with DU's stance by making reference to her utterances (*dediğiniz gibi/as you said*) and offering a (re)formulation (Heritage & Watson, 1979) of her prior account in lines 42 and 43. As it has been claimed by DU, SA indicates that in the other paragraphs there is information supporting the option in question while directing her gaze towards DU. DU then shifts her gaze to SA and nods, acknowledging SA's claim through a partial modified repeat (Kim, 2002; Stivers, 2005) of SA's utterance (*var/there is*) upon its completion in line 45. In an overlapping turn, GE also provides a (re)formulation (Heritage & Watson, 1979) in support of DU's account in lines 44 and 46 by specifying the paragraph number (*beşinci paragrafta/in the fifth paragraph*) where such information is included and locating the evidence for the correctness of the option in the text. GE proceeds with her turn by referring to the relevant information in the fifth paragraph with the adjective "*hızlı/fast*" but does not complete her unit of talk. In line 47, SA overlaps and shows agreement through the compliance token "*aynen/exactly*" and the use of an elongated acknowledgement token while directing her gaze towards GE.

Upon the other participants' display of agreement in the SPP that has been made relevant next by DU's base FPP of problematizing the item, DU takes the turn again in line 48 as all the co-participants are looking at their exam papers and initiates a non-minimal post-expansion sequence, which invites further elaboration and account of the problem (Stivers, 2013). Thus, in this post-expansion sequence, DU provides further elaboration on her previous account by explaining that the question is asked in general terms and reading aloud the stem, which SA agrees with in line 49 (*evet/yes*). As DU is reading aloud the stem, she elongates the first sound of the verb "*u:sed*" with a rising intonation and once again translates it into Turkish to emphasize the generalization in the question. Before DU finishes reading the stem, GE and SA in overlapping turns show stronger agreement in lines 51 and

52 by collaboratively (re)formulating the previously stated problem and providing an account in support of DU's explanation. In doing this, in line 51, GE orients to the violated testing principle first instantiated by DU by explaining from the students' perspective that they can also choose this option as a correct answer and demonstrates her understanding of the problem. On the other hand, in line 52, SA confirms DU by first acknowledging her account through the stance markers "doğru/*right*" and "tabi/*of course*" and then clarifying GE's (re)formulation of the problem by adding a condition to it (öğrenci sırasıyla gitmeyip başka bi yerden çıkarcak olsa/*if the student does not read in an order but derive it from somewhere else*), which GE further agrees with via a compliance token in line 55. Eventually, both GE and SA affiliate with the stance conveyed by DU.

Extract 4a demonstrates the act of problematizing an item in a language test IR interaction in a similar fashion as in the case of previously analyzed extracts. More specifically, this extract once again reveals that problematizing a language test item in an interaction with the tester and the other teachers as reviewers is not simply about stating the problem but as the problematizer DU has followed, it involves some steps in announcing the problem, proving it and reaching a consensus among the group so that the item can be improved. Firstly, DU's initiation of a pre-expansion sequence before stating the problem in the base sequence has enabled her to check and establish the necessary pre-conditions to proceed to the base sequence (Stivers, 2013). By eliciting confirmation of the answer that she chose in this pre-expansion sequence, DU ensures shared understanding and common ground on which to base the problem. In this way, she learns what the expected answer is so that she can proceed and problematize the other possible correct answer. Thus, the recipients' confirming responses act as a go-ahead response for problematizing the item and orienting to the violated testing principle in lines 15, 16 and 19 (Schegloff, 2007). Secondly, upon the announcement of the problem, DU goes through a continuous process of providing account and evidence from various perspectives and convincing everyone in the group of the problem. For instance, she completes her account in three stages throughout the interaction: first, she provides evidence from the reading text, secondly, she draws the attention to the way the stem is structured and makes a connection between the stem and the information presented in the text, and thirdly, she focuses on the generalized nature of the stem. She also continues her account even after everyone agrees on the problem by initiating a non-minimal post-expansion sequence. This could be because her aim is not just to be agreed with by

others but to make sure that they all understand the problem clearly in the same way so that they know where to make the correction in the next phase. In other words, they try to achieve mutual understanding about the problem.

As in the other extracts, extract 4a also illustrates that problematizing an item involves a dynamic process of collaborative inquiry. It is seen that not only the problematizer but also all the other participants actively and critically think about the item as prompted by DU's talk and attempt to make a contribution to the process of problematizing the item. This is evident first in GE's presentation of evidence from the reading text in response to DU's confirmation request for the correct option and then playing the devil's advocate and questioning the appropriacy of the key word "*efficient*" in the reading text from the students' and teaching perspectives by making reference to a shared past instructional event and textbooks that were studied previously (Can Daşkın & Hatipoğlu, 2019; You, 2015). She, thus, goes beyond simply confirming DU through an agreement token and carefully considers all the possibilities that could pose a threat to the validity of the expected answer.

Secondly, although the other co-participants just show listenership and agreement in the first two rounds of DU's account-giving, once DU completes her talk and makes her final comment about the problem in line 37, SA and GE take the floor and demonstrate their understanding of the problem by (re)formulating the problem previously stated by DU and providing further explanation from different perspectives (e.g., information in the reading text, students' approach to the item). SA, for instance, offers a (re)formulation (Heritage & Watson, 1979) of DU's account of the problem in lines 42 and 43 by orienting to the other paragraphs in the text and GE does so by specifying the paragraph number where the information in the option is available (i.e., paragraph five) and locating the evidence already presented by the problematizer in the text, which leads to complete agreement on the part of all the participants. Similarly, through the end of the post-expansion sequence in the interaction, GE and SA collaboratively provide an account in support of DU's explanation in lines 48 and 49 as GE (re)formulates the problem from the students' perspective by saying that the students can also go for the option in question and SA clarifies it by considering the possibility that the students might not read the text in an order. Thus, through those ranges of subsequent (re)formulations, the participants collaboratively explicate the problem initially raised by DU.



Thirdly, following her “pro forma” agreement (Schegloff, 2007), the tester SA’s projection of disagreement by presenting contradictory evidence from the reading text in line 38 leads to further thought on the problematic option. Eventually, it prompts even the tester to back down from her claim and disprove herself by considering her argument in line with DU’s account and thus, to realize the problem and agree with DU. Upon her challenging of the problem, the others also make an attempt to provide more justification. This also shows that in such task-oriented workplace interactions, contradictory view is not regarded as a threat to but as a source of validity and a way to resolve conflict (Davison, 2004). As Fulcher and Davidson (2007) indicate, during the validation process, there will be disagreements and alternative explanations that challenge the validity argument and test validation is about investigating those alternatives and successfully eliminating the counter-arguments to explain the evidence and support the use of a test for a given purpose (Fulcher & Davidson, 2007; So, 2014). Thus, such a collaborative way of understanding the problem and considering all the parameters of the problem in and through talk leads to its acceptance “in open discussion, through a process of dialogue and disagreement” (Fulcher & Davidson, 2007, p.21). Furthermore, the involvement of all the co-participants in the process of problematizing the item indicates that they all show ownership and feel responsible for the test, the outcome of which is important for all of them. In such a goal-oriented interaction, because the product is shared and affects everyone in the group, the problematizing, which is, by its very nature, a negative or a dispreferred act, is found to be carried out collaboratively and no defensiveness is involved on the part of the tester as the item writer as long as there is clear evidence for the problem.

The engagement of all the co-participants in the task of problematizing the item is also evident in their gaze behaviour. In an ordinary conversation, a disengaged gaze would make it sanctionable (Goodwin, 1981; Rossano, 2013). However, in this task-oriented interaction, the participants continuously look at their exam papers only sometimes directing their gazes at each other for a short time but this does not indicate disengagement; on the contrary, it signals that the participants are all engaged in the task and actively think about the matter at hand as they read the relevant part in the text, scan through it in search of evidence and go between the text and the item. In fact, this shows that they take the problem seriously and help each other identify it interactionally. As Goodwin (1981) acknowledges, engagement in a competing activity and looking at an object used for the accomplishment of such an activity is not a sign of diminished engagement in the conversation.

Regarding the institutional role distribution of the participants in this extract, there are indications of reviewer and tester roles and responsibilities in the IR interaction as in the case of other extracts. To begin with, upon DU's confirmation request for the correct answer, the teacher GE frames her confirmation from the students' and teaching perspectives (*kelimesinden çıkartıyorlar/they derive it from the word, bilirler/they would know, kitaplarda geçti/appeared in books*), while SA as the tester and writer of the item presents it from the perspective of test construction (*denizliyi hocam hedeflemiştim/I aimed for "d" miss*). In addition, DU directs the problem to the tester in particular and thus, orients to the tester's right, responsibility and obligation (Heritage, 2013) to learn about the problem as the writer of the item. This is most apparent in the reviewer DU's gaze direction towards the tester SA while explaining the problem (lines 24 and 33) as well as her sustained gaze at the tester and use of an endearment address term for her (*canım/dear*) in turn-final position while making a final remark about the problem (line 37).

Although the participants have a particular role in the IR interaction (e.g., reviewers, testers, problematizers of an item), due to the task-oriented nature of this interaction, they happen to put themselves in equal epistemic position depending on the kind of the task and problem that they focus on (Heritage, 2013). In this extract, since the problem is related to the options which derive from a reading text and since the participants all have equal epistemic access to the domain of information presented in this text, they treat the information as assumed and shared knowledge (Heritage, 2013). This is evident in DU's use of "*ya/you know*" in lines 25 and 48, SA's use of the stance marker "*tabi/of course*" as marked confirmation (Stivers, 2011) in line 52 as well as their constant orientation to the reading text as one is talking. This could be the reason why they can all make a contribution to the process of problematizing the item no matter what their role is in the IR interaction.

In extract 5a, the two teachers MI and AV teaching at B1 plus level review an item in a use of English quiz prepared for B1 plus groups and provide their feedback to the tester ME. They focus on a grammar item in a multiple-choice cloze test which tests students' knowledge of the past perfect tense. Students are expected to circle the verb in past perfect tense (i.e., *had seen*) among four different tenses (i.e., *was seeing, saw, had seen, has seen*) considering the context that is given in a text. However, the reviewer MI problematizes the

use of the word “*before*” in the target sentence in the text as a clue for past perfect tense as she believes that it does not fit well in the sentence and is inappropriate within the given context, making the meaning unclear. This is met with the other participants AV and ME’s agreement.

#### Extract 5a. Before, Segment a

- 1 MI: <ikide> benim şöyle bi  
 + all look at the item on their exam papers  
 + points at the word “before” in the item with a pencil in her hand
- 2 <sup>1#</sup> yorumum vardı burdaki (.) hani (.)  
*for two I have a comment such as this, in here well*  
 + ME turns gaze to MI’s paper and then nods  
 + AV scans through the target sentence with a pencil in her hand
- 3 ↑ had seen di mi [doğru] cevap  
*“had seen” is the correct answer isn’t it*  
 + shifts gaze to ME’s paper  
 + points at “had seen” on her paper with a pencil
- 4 ME: [evet ]  
*yes*
- 5 AV: ° huh-[huh °
- 6 ME: [o before’ [u ekledim [mecburen  
*I had to add that “before”*  
 + MI points at “before” on her paper with a pencil
- 7 MI: [er
- 8 [↑huh
- 9 beefore’u [eklemiştir  
*you have added*  
 + ME smiles
- 10 AV: [huh
- 11 (0.3)
- 12 MI: ya biz belki bildiğimiz için bize öyle geldi  
*you see, maybe it is because we know it that it*  
*seemed that way to us*  
 + shifts gaze to ME + AV underlines “before” on her paper
- 13 [bilemiyorum]  
*I don’t know*
- 14 ME: [çok mu net ] diceksiniz=  
*will you say that it is too obvious*
- 15 MI: =hayır çok- er hayır çok net değil  
*no too no it is not too obvious*  
 + shakes head
- 16 (0.3)

17 tam- (.) hiç öyle düşünmedim (.) [er  
**exact- I have never thought that way**  
+ shifts gaze to ME

18 ME: [°tamam°  
**okay**

19 MI: tam tersi  
**it is just the opposite**

20 (0.3)

21 ha böyle birden bire burda before fne<sup>#1</sup> geziyo diye  
+ bends forward and takes a close look at the paper

22 düşündüm  
**you know like all of a sudden I thought what**  
**"before" is doing here**  
#1 (lines 2-21) ME sustains her gaze at MI's paper and AV looks at her paper

23 (0.3)

24 ME: 2# şimdi ben şöyle düşündüm hocam (.) [aslında burda=  
**now this is what I thought miss actually here**

25 MI: [°uh-huh°

26 ME: =burda before yok orjinalinde  
**here there is no "before" in the original copy**

27 MI: huh-huh

28 ME: ama öğrenci  
**but the student**

29 (0.3)

30 had seen'e (0.2) [er gider- gidebilmesi için ]  
**goes for- to be able to go for "had seen"**

31 MI: [gidebilmesi için bişey lazım]  
**to be able to go for, something is required**

32 ME: bişey gerekiyo  
**something is required**

33 (0.2)

34 MI: huh-[huh

35 ME: [er: (.) o yüzden de<sup>#2</sup> before diye düşün[düm ama]  
**therefore I thought of "before" but**  
+ shifts gaze to AV

#2 (lines 24-35) all look at their papers

36 AV: [evet hh]  
**yes**

37 (0.4)

38 ME: öneriniz varsa değiştirebiliriz  
**we can change it if you have a suggestion**  
+ MI and AV scan through the target sentence with a pencil in their hands

39 (0.5)  
 40 MI: °er:°  
 41 (0.5)  
 42 AV: er bi bişey olması lazım  
           ***there should be something***  
           + turns gaze to ME's paper

As all the co-participants are looking at their exam papers, extract 5a begins with MI's pointing at the word "*before*" given in the sentence in the second item of the multiple-choice cloze test and her use of a preface (>ikide< benim şöyle bi yorumum vardı/*for two I have a comment such as this*) which projects the announcement of a problem in the item and draws others' attention to it as ME turns her gaze to MI's paper and AV scans through the target sentence with a pencil in her hand. However, before implementing the base sequence of problematizing the item, MI initiates a pre-expansion sequence in line 3 where she seeks confirmation for the answer that she chose (i.e., ↑had seen) by shifting her gaze to ME's paper and pointing at the option with "*had seen*" on her paper with a pencil. Then she receives a confirming response from both the tester ME and the other reviewer AV in lines 4 and 5.

ME proceeds with her turn in order to elaborate on the way the item is constructed in line 6 by anticipating and pre-empting the projected problem, that is the use of "*before*" in the target sentence, which is prompted by MI's pointing at "*before*" on her exam paper with a pencil and ME's sustained gaze at it. Thus, ME ends up articulating the base FPP of problematizing the item indirectly in the form of a pre-emptive account for the problem (o before'u ekledim mecburen/*I had to add that "before"*) by reference to "*before*" with the recognitional demonstrative "o/*that*". This is accepted by MI through her use of an acknowledgement token with a rising intonation in line 8 and then her full modified repeat (Stivers, 2005) of ME's explanation with a smiley voice accompanied by ME's smile in line 9, which signal that ME was right in her anticipation of the cause of the problem.

In line 10, in overlap with MI's turn, AV displays understanding of the matter at hand via an acknowledgement token and then orients to the target sentence with a pencil in her hand, underlining "*before*" on her paper as MI retains the floor to state the problem more explicitly after a brief gap of silence. In doing this, by shifting her gaze to ME, MI uses a preface

where she downgrades her upcoming claim for the problem through the hedges “belki/*maybe*” and “bilemiyorum/*I don’t know*” and the account that their realization and shared knowledge that “*before*” was added to the sentence might have caused them to see it as a problem. Then before MI states the actual problem, with her gaze sustained at MI’s paper in line 14, ME takes over in an overlapping turn and once again anticipating and pre-empting the projected claim with regard to the problem in the item, she offers a candidate evaluation of the use of “*before*” by voicing MI’s next unit of talk (diceksiniz/*you will say*) and requesting MI’s confirmation for it. Without any time lapse, MI provides an immediate response that is non-confirming. As MI is trying to formulate her response, she offers the no response (hayır/*no*) two times followed by a repetition of ME’s prior evaluation and then a claim of its non-acceptance (değil/*not*) along with a disagreeing head shake. After a gap of brief silence, MI shifts her gaze to ME and displays a more overt disagreement with her assertion in line 17, which is received with a minimal expansion acceptance token “°tamam°/*okay*” by ME. Between lines 19 and 22, MI implements the actual act of problematizing the item by first projecting it as contrary to ME’s candidate evaluation (tam tersi/*it is just the opposite*) and then with a smiley voice, explaining the problem that the word “*before*” does not fit well in the target sentence and that it stands out by bending forward and taking a closer look at her paper.

Following MI’s announcement of the problem, ME provides an explanation that justifies MI’s assumption and claim with regard to the use of “*before*” in the target sentence. In doing this, ME once again elaborates on the way she constructed the item and why she had to use “*before*” between lines 24 and 35 by this time offering more details while all the co-participants’ gaze is at their own exam papers. She makes it clear that in the original version of the text, “*before*” was not used but because something was required for the students to be able to choose the past perfect tense, she thought to insert “*before*” into the sentence in question. As ME is making this explanation, MI first shows listenership in line 25, and then displays understanding via the use of an acknowledgement token in line 27. Following the pause in ME’s as-yet-incomplete account when she has just mentioned the correct option “*had seen*” with an emphasis on the verb, MI uses anticipatory completion (Lerner, 1996a) in line 31, uttering the explanation ME has been trying to formulate and displaying her uptake of ME’s item writing process and thus, affiliating with her (Hayashi, 2013). In an

overlapping turn with MI's anticipatory completion, ME also continues her account with a hesitation marker and a self-initiated self-repair, cutting off the verb "gider/*goes*", substituting it with MI's utterance (*gidebilmesi için/to be able to go*) and then employing a full modified confirming repeat (*bişey gerekiyo/something is required*) (Kim, 2002; Stivers, 2005) of MI's anticipatory completion (*bişey lazım/something is required*), which MI agrees with in line 34. ME, then, makes her final remark in line 35 by shifting her gaze to the other reviewer AV and concluding that this is the reason why she thought of using "*before*" in the sentence. This is met with a responsive uptake by AV, who shows agreement through the compliance token "*evet/yes*". After a gap of silence, ME then accepts to make a change in the item if they have a suggestion, which conveys her agreement with the problem stated by MI. Through the end of her utterance, MI and AV start to scan through the target sentence with a pencil in their hands and silence ensues as they are all engaged in the test material. Finally, by turning her gaze to ME's paper, AV asserts that something is required in the target sentence, which makes the phase of SC relevant next.

Extract 5a demonstrates how a problem in a language test item is collaboratively identified and explained by means of some interactional resources. The reviewer MI and the tester ME are the active agents who play an important role in the process of problematizing the item. MI as the 'problematizer' of the item does not simply state the problem but uses some interactional features that lead to the realization, collaborative explanation and acceptance of the problem in and through the IR interaction. For instance, MI employs a pre-expansion sequence seeking confirmation for the correct answer (line 3) as in extract 4a, which serves the base sequence of problematizing the item by establishing the necessary pre-conditions and laying the groundwork (Stivers, 2013). This is because she needs to ensure what the expected answer is and that there is no other problem that she could not foresee so that she can proceed to the base sequence. Furthermore, MI's use of a preface two times during the interaction before announcing the problem enables the tester ME to anticipate the projected problem and engage in a pre-emptive act of problematizing the item, which leads to further negotiation and evaluation of the item from different perspectives. MI's first preface in the beginning of the sequence (lines 1-2) is accompanied by her non-verbal behaviour of pointing at the word "*before*" on her paper with a pencil, which projects the announcement of a problem in the use of "*before*" in the item, drawing others' attention to it. As a result,

ME shifts her gaze to MI's paper, anticipates that the problem is about the use of "*before*" in the target sentence and eventually, provides a pre-emptive account for the projected problem by reflecting on the way she constructed the item. In a way, she makes a confession early on without waiting for others to ask if "*before*" was added by her and having to respond to their questions. This also saves MI's time for problematizing the item by confirming in advance her prediction about ME's insertion of "*before*" into the sentence as correct (before' u eklemiştir/you have added). As for MI's second preface before stating the problem more explicitly (lines 12-13), it also paves the way for ME's anticipating and pre-empting the projected claim with regard to the problem in the item in the form of a candidate evaluation of the use of "*before*". However, this time her pre-emption is rejected by MI, who shows disagreement and then provides her own evaluation of the use of "*before*" in the item, at which point the identified problem becomes apparent. Although ME was right in her anticipation of the cause of the problem in response to the first preface, her pre-emptive claim following the second preface with regard to why "*before*" would cause a problem did not match MI's stance about the problem. Such a process allows the item to be evaluated jointly from different perspectives through negotiation, which then enables the tester ME to account for MI's claim by elaborating on the way she constructed the item and to eventually reach a consensus on the problem. ME provides the necessary background information about why she had to use "*before*" in the target sentence and only when she ensures shared understanding among the group as signalled by MI's anticipatory completion (Lerner, 1996a) and AV's compliance token, she displays a stance of acceptance of the problem by proposing a change in the item.

As in the other extracts, the participants' gaze behaviour once again illustrates that all the participants are actively engaged in the task of problematizing the item. It is seen that throughout the interaction, the participants continuously look at their exam papers only sometimes directing their gazes at each other for a short time but this does not indicate disengagement; on the contrary, it signals that the participants actively think about the matter at hand in such a competing activity (Goodwin, 1981; Rossano, 2013). Although the reviewer AV is relatively more passive, her orientation to the test material and her sustained gaze at her paper throughout the extract, scanning through the target sentence with a pencil in her hand at certain points and underlining "*before*" on her paper while the problem is presented all show her engagement in the task, her constant thinking about the issue as well as her understanding of the problem.



With respect to the institutional role distribution of the co-participants in this extract, there are indications of reviewer and tester roles and responsibilities in the IR interaction as in the case of other extracts. To begin with, in response to ME's pre-emptive account for the projected problem, MI's use of an acknowledgement token with a rising intonation in line 8, her modified repeat (Stivers, 2005) of ME's assertion in line 9 and producing this with a smiley voice all signal that MI has already predicted and realized that "*before*" was added by the tester. In this way, because she cannot go first in sharing her prediction (*beɪfoʊə' u ekleɪʃsɪnɪ/you have added*) when ME takes over and pre-empts the projected problem, she claims her epistemic right to confirm ME and to assert epistemic authority and control, orienting to her role as the reviewer who has done her job of reviewing and evaluating the item and her right to present the problem which she previously identified (Heritage & Raymond, 2005; Stivers, 2005). On the other hand, before announcing the actual problem, MI produces a preface that downgrades her upcoming claim for the problem through hedges and an account in lines 12 and 13. This could be because the problem is directly related to the tester's practice in the construction of the item as revealed by the tester herself in the prior turns. Based on this, as the first assessor of the item, MI might not want to implicate epistemic priority about a matter that directly concerns the tester's practice (Heritage & Raymond, 2005). In fact, this could be the reason why the tester attempts to anticipate and articulate the problem pre-emptively before MI does once she realizes that the problem is about her use of "*before*" in the item. Additionally, as in extract 3a, while providing the preface, MI also uses an inclusive language for the reviewers via the emphasized pronoun "*bɪz/we*" and she shares the responsibility for problematizing the item with the other reviewer, establishing common ground on which to base the problem. In this way, she indicates that not only herself but also the other reviewer AV is aware that "*before*" was added to the sentence by the tester and they adopt the same stance toward the item, orienting to their shared epistemic knowledge and their equal epistemic access to the matter at hand (Heritage, 2013).

With respect to the tester roles, in the pre-expansion sequence, MI directs her confirmation request to the tester ME by shifting her gaze to ME's paper, which signals ME's right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item as the writer and thus, the 'owner' of the item. Similarly, through her explanation of the item writing process, ME orients to her role as the tester who has first-hand knowledge (Type 1 knowledge, Pomerantz, 1980) of the issue that has caused a problem in the item

(i.e., the use of “*before*” in the item) and thus, asserts epistemic priority over the matter. For the same reason, when MI employs an anticipatory completion during ME’s explanation of the way she constructed the item, ME uses a modified confirming repeat of MI’s anticipatory completion, indicating her epistemic right and responsibility to explain the item writing process which she has first-hand access to (Heritage & Raymond, 2005; Stivers, 2005).

In extract 6a, the two teachers DU and GE teaching at B1 level review an item in a note-taking listening quiz prepared for B1 groups and provide their feedback to the tester SA. The listening section in this quiz aims to test the students’ note-taking skills and thus, they are to first take notes while listening to a lecture on phobias and then fill in an outline according to the notes they have taken. In this extract, the reviewers focus on one of the blanks in the given outline where the expected answer is “*chemicals*”. However, GE problematizes this item by claiming that “*physical reactions*” is also possible as a correct answer and if the students happen to write it in the blank in question, they might have to accept it as correct. In the end, the other participants DU and SA also agree with the presented problem.

#### Extract 6a. Physical Reaction, Segment a

- 1     DU: °öbüründe chemicals°  
       *in the other it is "chemicals"*  
       + all look at the item on their exam papers  
       + DU points at the item with a pencil in her hand  
       + SA shifts gaze to DU’s paper
- 2     SA: huh-huh
- 3     GE: chemicals
- 4       =physical reaction<sup>1#</sup>derse gerçi onu yazamaz (.)  
       *if s/he says "physical reaction", though s/he*  
       *cannot spell it*  
       + DU shifts gaze to SA’s paper
- 5       tabi  
       *of course*
- 6       (0.5)
- 7       make physical reactions  
       + turns gaze to SA
- 8       (0.5)
- 9       onu da kabul eder miyiz?  
       *would we accept that too?*  
       + turns gaze to DU then SA then back to her paper
- 10      (0.6)

- 11 SA: uh:m<sup>#1</sup>  
+ GE shifts gaze to SA  
#1 (lines 4-11) DU and SA sustain gaze at their papers and DU changes her gaze repeatedly between the note-taking sheet and the exam paper
- 12 GE: <sup>2#</sup>çünkü beyin (.) fiziksel reaksiyonlar yapıyo  
**because the brain makes physical reactions**  
+ SA shifts gaze to GE
- 13 (0.6)  
14 beyinde [fi- fi]ziksel  
**in the brain physical**
- 15 SA: [evet ]  
yes
- 16 GE: reak[siyonlar >oluyo sanırım öyle bişey<]<sup>#2</sup>  
**reactions occur, I think something like that**  
#2 (lines 12-16) GE sustains her gaze at SA and DU looks at the item and then changes her gaze between her notes on the note-taking sheet and the item on the exam paper
- 17 SA: [bi (.) scripti alıyım mı? ] hocam=  
**shall I take the script miss**  
+ moves towards the computer to print out a copy of the script
- 18 GE: =ya da chemicals are making er (0.3) er er ki-  
**or**  
+ shifts gaze to DU  
+ DU shifts gaze to GE and then nods
- 19 physical reactions o- orasını [yakalayamadım]  
**I couldn't catch that bit**  
+ GE and DU shift gaze to the computer screen  
+ turns gaze to DU
- 20 DU: [hiç er ]  
**any**
- 21 hiç er: hata olamaz o şekilde  
**there cannot be any mistake that way**  
+ turns gaze to GE
- 22 =yani (0.2) [hata derken er doğru]  
**I mean when I say mistake correct**  
+ shifts gaze to SA
- 23 GE: [physical reaction- ]  
+ shifts gaze to SA
- 24 DU: <sup>3#</sup>[cevap ver- kab-  
**answer give-**
- 25 SA: [yanlış <ol(.)maz>  
**it won't be wrong**
- 26 DU: yanlış olmaz=  
**it won't be wrong**
- 27 SA: =[huh-huh ]
- 28 DU: =[yani doğru] cevap kabul edilmesi gerekir orda.

- I mean it should be accepted as a correct answer there*
- 29 SA: tamam hocam ben (.) bi<sup>#3</sup> (.) script'i alıyım  
*okay miss let me get the script*  
 #3 (lines 24-29) GE and DU sustain their gaze at SA
- 30 =orda aslında bişey vardı [da ben]  
*there was actually something there but I*
- 31 DU: [yani ]  
*you know*  
 + shifts gaze to the computer screen
- 32 şöyle (.) [anlamı verecek  
*will give the meaning as such*
- 33 SA: [uzatmadım  
*didn't extend it*
- 34 GE: [bişey vardı  
*there was something*  
 + turns gaze to DU
- 35 DU: şekilde (.)<sup>4#</sup> bir sözcükle ifade edebilir >şimdi<  
*s/he can express it with a word in a way that, now*  
 (0.6)
- 36 lecture'ın (.) o (.) kıs[mında  
*in that part of the lecture*
- 38 SA: [ °hu:h °
- 39 DU: diyo işte  
*it says that*  
 + turns to her notes on the note-taking sheet
- 40 (0.8)
- 41 er
- 42 (0.4)
- 43 <sup>5#</sup> işte daha çok (0.2) vücut daha (.) işte terler er:  
*that much more, that the body sweats more*
- 44 ateş (.) iş:- (0.2) [şey yapar]  
*heat does thing*
- 45 GE: [basar ]<sup>#4</sup>  
*surrounds your body*  
 #4 (lines 35-45) GE looks at the item on her paper and SA takes the printout
- 46 kalbi a[tar  
*one's heart beats*
- 47 SA: [huh-huh
- 48 DU: [kalp atmaya başlar (.) er:  
*the heart starts beating*
- 49 (0.7)

- 50 görmesi, koku alması: (0.2) er: <sup>#5</sup> değişir  
**the way one sees and smells changes**  
+ GE shifts gaze to DU + shifts gaze to GE  
#5 (lines 43-50) DU reads from her notes
- 51 GE: evet  
**yes**
- 52 DU: farklı algı[lar=  
**one perceives differently**  
+ shifts gaze to SA
- 53 GE: [↑evet  
**yes**
- 54 DU: =hatta a- (ağ-) çok ağır bişeyi (.) kolayca  
55 [kaldırabilir gibi ]  
**in fact one can easily lift something too heavy and so on**  
+ moves arms upwards with gaze at SA then at her paper
- 56 GE: [fiziksel reaksi ]yonlar evet.  
**physical reactions yes**  
+ turns gaze to her notes on the note-taking sheet then the item on her exam paper
- 57 DU: öyle de ↑diyebilir reactions  
**one can also say that**  
+ shifts gaze to SA
- 58 SA: evet hocam (.) zaten aslında öyle bir (0.2) er: ifade  
+ comes back to the desk with the printout of the script in her hand  
+ gaze on the script  
+ DU turns gaze to the script
- 59 var orda söyle diyo  
**yes miss actually there is already such an expression there, it says**
- 60 (0.5)  
61 your brain quickly makes chemicals that cause  
+ reads aloud the relevant sentence on the script with a pencil in her hand  
+ DU follows SA
- 62 different physical reactions [diyo: ]  
**it says**  
+ shifts gaze to GE
- 63 DU: [huh-huh]
- 64 GE: °huh-huh°

Extract 6a begins with DU's giving the correct answer "*chemicals*" by pointing at the item with a pencil in her hand as all the co-participants are looking at their exam papers. Just as DU utters the correct answer, SA shifts her gaze to DU's paper and then acknowledges the answer with "huh-huh". In line 3, GE first repeats the expected answer, confirming DU and then rushes to launch the base FPP of problematizing the item between lines 4 and 9 by

asking whether they would also accept “*physical reactions*” as a possible correct answer. She, thus, challenges the item by presenting the problem as a question for everyone in the group to think about using the inclusive “*we*” pronoun (kabul eder miyiz?/*would we accept*). In doing this, she also uses a parenthetical insert where she asserts that the students cannot spell “*physical reactions*” correctly as an additional note to what she was saying with regard to “*physical reactions*”. Additionally, upon GE’s repetition of the correct answer in turn-initial position, DU shifts her gaze to SA’s paper and while stating the problem following the parenthetical insert, GE orients to the other participants by shifting her gaze from her exam paper to SA in line 7 and then to DU and SA again respectively in line 9, and turns back to her paper just prior to the completion point of her turn. At the same time, SA sustains her gaze at her exam paper and DU changes her gaze repeatedly between her note-taking sheet and the exam paper.

After a gap of silence, SA takes a turn with a turn-initial elongated hesitation marker, at which point GE shifts her gaze to SA again. In line 12, GE takes over and provides an account for the problem (çünkü beyin (.) fiziksel reaksiyonlar yapıyo/*because the brain makes physical reactions*), which elicits SA’s gaze. She, then, rephrases the account in line 14 and downgrades her claim through an expression of uncertainty (sanırım öyle bişey/*I think something like that*) uttered in a fast pace. During the rephrasing, SA overlaps in line 15 and displays listenership via the continuer “*evet/yes*” while DU continues to change her gaze between her notes on the note-taking sheet and the item on her exam paper.

Upon GE’s presentation of an account with less certainty, SA moves towards the computer to print out a copy of the tapescript. Meanwhile, GE continues her turn of account-giving in lines 18 and 19 with an alternative possible explanation by trying to remember what was said exactly in the lecture as evident in her use of hesitation markers, an intra-turn pause and her statement that she could not catch the part about “*physical reactions*” in the lecture. During her turn-at-talk, GE and DU turn their gazes to each other followed immediately by DU’s nodding and then to the computer screen as GE displays hesitation about what was said in the lecture. Just prior to the completion of her turn, GE shifts her gaze to DU who then overlaps in line 20 to respond to the question GE previously asked in the base FPP of problematizing the item by turning her gaze to GE in line 21. DU produces a non-

confirming (Raymond, 2003) elaborative response where she indirectly supports and endorses GE's stance with regard to the violated testing principle by acknowledging that it will not be wrong to say "*physical reactions*" and thus, (re)formulating (Heritage & Watson, 1979) the problem from a different perspective (*hiç er: hata olamaz o şekilde/there cannot be any mistake that way*), which departs from the terms of GE's polar question (*onu da kabul eder miyiz?/would we accept that too?*) (Raymond, 2003). In line 22, DU quickly attempts to rephrase and (re)formulate her just produced remark to make her stance clearer but just as DU turns to SA sitting in front of the computer, SA overlaps in line 25 to complete DU's ongoing turn by using an anticipatory completion (Lerner, 1996a). SA utters the explanation that DU has been searching for as evident in DU's intra-turn pause, her use of a hesitation marker and self-initiated self-repair (*cevap ver- kab-/answer give-*), and displays her understanding of the problem by offering an interpretation of what DU has just said in line 21. Additionally, at one point in DU's turn-in-progress, GE overlaps in line 23, repeating the second possible correct answer "*physical reaction*" by shifting her gaze to SA but she does not continue and cuts off her utterance when SA takes over for an anticipatory completion. As GE and DU sustain their gazes at SA, DU cuts off her utterance in line 24 and acknowledges SA with a responsive uptake by fully repeating her anticipatory completion in line 26, which SA further agrees with via an acknowledgement token. In line 28, DU rushes through into another unit of talk where she repeats her prior self-repaired cut-off utterance and then completes it, eventually producing a response that is now in line with the terms, constraints and presuppositions of GE's polar question (*kabul edilmesi gerekir/it should be accepted*) and thus, displays alignment with GE. This is received with SA's use of a minimal expansion acceptance token "*tamam/okay*" followed by the address term "*hocam/miss*" and her offer to get a copy of the tapescript.

Once GE's polar question is answered with a confirming response and the problem announced and explained in the base FPP is agreed with by all the co-participants, SA initiates a non-minimal post-expansion sequence in line 30 by orienting to the tapescript, reflecting on her item writing process and explaining that there was something in the tapescript which she did not include in the item in order not to make it long. She, thus, provides further elaboration in support of GE's prior account with respect to the existence of a relevant piece of information in the script, which was presented with less certainty. During

SA's turn-at-talk, DU overlaps in line 31 and by shifting her gaze to the computer screen, she provides an extension to her prior turn, adding that the students can express it with a word that will give the meaning as such. In overlap with SA's final word of her unit of talk, GE also displays agreement and alignment in line 34 by using a partial modified repeat (Kim, 2002; Stivers, 2005) of SA's utterance, repeating the part that is relevant to the claim she offered earlier (i.e., *bişey vardı/there was something*) and just prior to the completion of her turn, she turns her gaze to DU, whose turn is still in progress. The post-expansion sequence, thus, proceeds with DU's continuation of her unit of talk already in progress. In line 37, as GE looks at the item on her paper and SA takes the printout, DU orients to the relevant part of the lecture, at which point SA shows listenership, and then DU turns to her notes on the note-taking sheet in line 39.

After a gap of silence, DU begins to provide concrete evidence by reading from her notes and listing the physical effects of fear one by one collaboratively with GE between lines 43 and 55. During her explanation of the first effect of fear in line 44, as she searches for a word, GE uses an anticipatory completion (Lerner, 1996a) in line 45 following the pause in DU's turn and continues to articulate the second effect in their list in line 46, which SA agrees with through an acknowledgement token (*huh-huh*). In line 48, DU immediately confirms GE through a full modified repeat (Kim, 2002; Stivers, 2005) of her utterance and after an elongated hesitation marker and a gap of silence, DU provides the third physical effect of fear, at which point GE shifts her gaze to DU. Just as DU directs her gaze at GE through the end of her utterance, GE displays agreement via the compliance token "*evet/yes*". Then in line 52, DU articulates the fourth effect that she noted by shifting her gaze to SA, with which GE once again agrees by producing the compliance token "*↑evet/yes*" with a rising intonation in line 53. With her gaze at SA, DU rushes to explain the final effect of fear in lines 54 and 55 by moving her arms upwards to demonstrate that one can lift something heavy easily.

Then in line 56, by turning her gaze to the notes on her note-taking sheet and then to the item on her exam paper, GE overlaps and relates what DU has been saying to the problem by once again repeating the second possible correct answer "*fiziksel reaksiyonlar/physical reactions*" and then acknowledging DU with "*evet/yes*". In line 57, shifting her gaze to SA, DU makes a final remark and



(re)formulates the problem by concluding that the students can also say “reactions” with a rising intonation on the first syllable of the verb “↑diyebilir/*one can say*”, which assures the students’ possible action of writing down “*physical reactions*” in the blank. This is met with SA’s agreement through the use of a compliance token as SA comes back to the desk with the printout of the script in her hand and as she and DU orient to the script with their gazes on it. Then, SA further supports and proves what DU and GE have said with regard to the violated testing principle based on concrete evidence by referring to the relevant sentence in the script in lines 58 and 59 and reading it aloud with a pencil in her hand in lines 61 and 62. Upon the completion of her reading aloud, SA shifts her gaze to GE and the extract ends with both GE and DU’s display of agreement respectively by using the acknowledgement token “huh-huh”.

Extract 6a demonstrates how reviewers together with the tester collaboratively identify, explain and prove a problem in a language test item in and through interaction. First of all, GE’s announcement of the problem through a question directed to everyone in the group with the use of the inclusive “we” pronoun paves the way for a collaborative search of evidence for the problem. Therefore, before answering GE’s question, SA and DU directly orient to the test material and start to think about the matter at hand. DU then responds to GE’s question only when she finds out supportive evidence by changing her gaze repeatedly between her notes and the exam paper accompanied by GE’s account-giving. Thus, DU’s delayed response is not indicative of an upcoming dispreferred or disconfirming response (Pomerantz, 1984a) but rather it signals the participants’ careful thought on and preliminary inquiry into the issue raised by GE in order to be able to provide a valid response. Eventually, with her delayed response, DU supports and endorses the stance adopted by GE with regard to the problem in the item.

Secondly, as part of the collaborative inquiry into GE’s question, a dynamic process of account-giving and evidence-seeking is followed. It begins with GE’s formulation of an account based on what was said in the lecture but it is provided with less certainty as she cannot exactly recall the part in the lecture about “*physical reactions*”, which invites further search for evidence for the problem. Following GE’s account-giving, DU presents more concrete evidence from her notes by listing the physical effects of fear through collaborative turn completions with GE and in the post-expansion sequence, SA provides the third round

of account-giving by reading aloud the sentence in the script which has been sought by the IR group and which presents first-hand evidence to prove the problem. As a result, the co-participants orient to three different sources of evidence based on what each participant has direct access to (i.e., the lecture that was listened to, notes taken on the note-taking sheet, a copy of the tapescript) and produce an account from different perspectives until the problem is well ensured. In this respect, SA's initiation of a non-minimal post-expansion sequence plays an important role in inviting further elaboration and account of the problem (Stivers, 2013). Although an agreement was already reached on the problem, through such a post-expansion sequence, the participants make sure that there is strong evidence for the problem so that they know where to make the correction in the next phase.

Thirdly, the problem is stated not just by the problematizer and is not simply accepted by the others but rather it is expressed by the co-participants in their own terms through (re)formulations (Heritage & Watson, 1979) and when stating the problem, they consider all the dimensions of the problem and critically evaluate it, which leads to a better understanding of the problem. For instance, when announcing the problem in the first place, GE uses a parenthetical insert where she asserts that the students cannot spell "*physical reactions*" correctly as an additional note to what she was saying with regard to "*physical reactions*". By analyzing the problem from the students' perspective (*yazamaz/s/he cannot spell it*), she carefully considers all the possibilities that might affect the acceptance of the problem by the group. In addition, when responding to GE's polar question, DU does not show agreement with the problem simply with a "yes" but instead, she (re)formulates the problem through a non-conforming elaborative response. As a result, her non-conforming response and her (re)formulations of her statement of the problem enable her to go beyond the terms and constraints of GE's question and thus, to evaluate the presented problem from different perspectives while at the same time supporting GE's stance and ensuring mutual understanding by later rephrasing her utterance to conform to the terms of GE's question (line 28). Based on this, SA's anticipatory completion of DU's response turn in line 25 leads to the collaborative statement of the problem. This once again brings out different viewpoints that convey the same stance towards the problem and demonstrates the participants' understanding of the matter at hand.

All the three participants' active engagement in problematizing the item is also evident in their gaze behaviour. Throughout the interaction, they orient to their own and each other's

test material to check their answers and find out about the expected answer, to their notes on the note-taking sheet, and to the tapescript first seen on the computer screen and then taken as a printout in search of evidence for the problem. It is seen that they direct their gazes at each other only for a short time at certain points during the interaction, such as when GE is announcing the problem. This shows that in such a competing activity, the participants are all actively engaged in the task and constantly think about the matter at hand in an attempt to help each other identify the problem interactionally and to make a contribution to the process of problematizing the item (Goodwin, 1981; Rossano, 2013).

With respect to the institutional role distribution of the participants in this extract, there are indications of reviewer and tester roles and responsibilities in the IR interaction as in the case of other extracts. To begin with, DU's non-conforming response to GE's question announcing the problem, her (re)formulations of the problem, her modified confirming repeat of GE's utterance in line 48 as well as her full repetition of SA's anticipatory completion in line 26 all "assert the primacy of the second speaker's rights to make the statement" and thus, claim her epistemic right to confirm the first speaker (Heritage & Raymond, 2005; Stivers, 2005, p.131). In this way, DU orients to her institutional role as the reviewer who has done her job of taking notes while listening to the lecture, answering the questions as a student would and evaluating the item just like GE has done. By going first in responding to GE's question, DU also orients to her epistemic responsibility and right to respond to GE as the reviewer who has gone through the same procedure as GE, listening to the lecture, taking notes and having a thought on the item. Based on this, DU implies epistemic authority in her response after scanning through her notes by stating the problem with greater certainty than GE in line 21, which leads to its immediate acceptance by the tester SA. Similarly, through her modified repeat of SA's assertion in line 34, GE asserts epistemic authority and control, orienting to her role as the reviewer and problematizer of the item who first realized that there was another possible correct answer in the lecture and thus, who first announced the problem (Heritage & Raymond, 2005; Stivers, 2005). As the problematizer of the item, GE also monitors the talk by repeating the second possible correct answer "*physical reactions*" twice in lines 23 and 56 during DU's turn-in-progress, and thus, relating DU's turn of account-giving to the problem and reminding the group of the answer that poses a problem. In addition, DU and GE's roles as reviewers and as the teachers of the group of students who are subject to the test are also evident in their statement and explanation of the problem from the students' perspectives as well.

Regarding the tester roles, SA indexes her right, responsibility and obligation to reflect on her item writing process in the post-expansion sequence as in extract 5a and to provide the tapescript by taking a printout of it, orienting to her role as the tester who has first-hand knowledge (Type 1 knowledge, Pomerantz, 1980) of the issue that has caused a problem in the item (i.e., the exclusion of a piece of information that is present in the script from the outline). It is also seen that throughout the interaction, the reviewers mostly shift to and sustain their gazes at the tester SA while stating and explaining the problem (lines 7, 9, 12-16, 22-29, 52-55, 57) and thus, orient to her right to learn about the problem as the ‘owner’ of the item so that the item she has prepared can be improved.

In extract 7a, the teacher CI teaching at B1 level reviews an item in a use of English quiz prepared for B1 groups and provides her feedback to the tester AY. They focus on a grammar item in a binary-choice test which tests students’ knowledge of the modal verbs they have studied in class. They are to circle the modal verb among two alternatives (i.e., *should*, *might*) that best fits in the blank and the context given in a text. The expected answer is “*should*”; however, the reviewer problematizes the item, claiming that the distractor “*might*” can also be possible as a correct answer. In the end, the group reaches a consensus on the problem.

#### Extract 7a. Should Might, Segment a

- 1 CI: burda da (.) sonuncusunda da should (.) olduğunu  
+ points at the item with a pencil in her hand + underlines “should”  
+ both look at the item on CI’s exam paper
- 2 dü[şündüm=  
**and here for the last one I thought it was “should”**
- 3 AY: [huh-huh
- 4 CI: =ama might riskli mi diye bi düşünüyorum  
**but I am thinking if “might” is risky**  
+ double underlines “might”
- 5 AY: huh-huh (.) bi bakıyım (.) .hhh  
**let me see**  
+ turns gaze to her exam paper
- 6 ° the crabs are °  
+ starts to read the relevant part on her exam paper  
+ produces thinking face and puts her hand on her forehead
- 7 ((AY silently reads the relevant part in the text and  
CI sustains her gaze on her exam paper for 13.9s))
- 8 CI: hani <sup>1#</sup>bu sebepten ötürü in[gilizler] (.)  
**well because of this reason the english**

- 9 AY: [huh-huh ]
- 10 CI: dikkatli (.) [olabilirler ve]  
**might be careful and**
- 11 AY: [olabilirler ]  
**might be**
- 12 CI: koruyabilirler<sup>#1</sup>  
**might protect**  
+ turns gaze to AY  
#1(lines 8-12) CI translates the sentence into Turkish by putting “might” in it and moves the pencil along the sentence, pointing at the verbs with “might” while AY looks at her exam paper
- 13 AY: huh-[huh ]
- 14 CI: [gibi] (.)  
**like that**  
+ shifts gaze back to her paper
- 15 AY: o da olabilir doğru  
**it can also be possible, right**
- 16 (1.2)
- 17 CI: oraya başka bişey ne olabilir  
**what else can be put there**  
+ produces thinking face and puts hand under her chin  
+ AY gets the textbook and searches in it

As CI and AY both look at the item on CI’s exam paper, extract 7a begins with the reviewer CI’s giving the correct answer “*should*” for the last item in the binary-choice test by pointing at the item with a pencil in her hand, putting an emphasis on “*should*” and then underlining it right after uttering it. Just prior to the first possible completion point of CI’s unit of talk, her answer is acknowledged with “huh-huh” by the tester AY in line 3. Then CI rushes to launch the base FPP of problematizing the item in line 4 with a *but*-prefaced turn which projects the announcement of a problem in the item. According to CI, the other option “*might*” can be risky as a possible correct answer. While articulating the problem and orienting to the violated testing principle, CI double underlines the risky option “*might*” and downgrades her claim through a mitigating embedded question as in extract 4a, which leaves it to AY to further think about the item and recognize the problem herself.

As a result, AY displays understanding of the issue raised by CI via an acknowledgement token in line 5 and then orients to the test material by shifting her gaze from CI’s exam paper to hers and asking to have a look at the text. In line 6, she starts to read the relevant part in the text silently by producing a thinking face with her hand on her forehead while CI sustains her gaze on her exam paper. Following a lapse of 13.9 seconds, as AY is looking at

her exam paper, CI takes over and provides an account for the problem between lines 8 and 12. In doing this, she translates the target sentence into Turkish by putting “*might*” in it and moves the pencil in her hand along the sentence, emphasizing the adjective “*dikkatli/careful*” and pointing at the verbs with “*might*” (i.e., *olabilirler/might be*, *koruyabilirler/might protect*). As CI initiates her turn of account-giving, AY overlaps and displays understanding with “huh-huh” in line 9. During CI’s turn-in-progress, AY also overlaps in line 11 and engages in choral co-production (Lerner, 2002) with the current speaker CI by co-producing one of the verbs in the target sentence transformed into a form with “*might*” in Turkish (i.e., *olabilirler/might be*). In this way, AY affiliates with the stance taken by CI in problematizing the item (Hayashi, 2013). CI ends her turn by directing her gaze at AY, who then shows agreement by using the acknowledgement token “huh-huh” at the transition-relevance place. Then CI shifts her gaze back to her paper and AY overtly confirms the problem by (re)formulating the problem initially stated by CI and acknowledging that “*might*” can also be possible as a correct answer in line 15. After a gap of silence, CI directly starts to think about an alternative option with a display of a thinking face by putting her hand under her chin, and AY gets the textbook and searches in it, which projects the phase of SC as the relevant next action in the IR interaction.

Extract 7a demonstrates how the reviewer’s way of stating and explaining the problem leads to its recognition and acceptance. To begin with, CI implements the base FPP of problematizing the item through an embedded question (*might riskli mi diye bi düşünüyorum/I am thinking if “might” is risky*), which leaves room for collaborative thought and inquiry on the raised issue. Eventually, AY directly orients to the test material in search of evidence in order to be able to confirm the problem and CI rechecks the matter at hand by keeping her gaze on the relevant part in the test. In this way, as the first assessor of the item, CI might not want to implicate epistemic priority (Heritage & Raymond, 2005) and might not want to sound too assertive about an issue which concerns the domain of information given in the text of the item and which both CI and AY have equal epistemic access to, but rather invite careful thought on the problem. As a result, only after having a thought on the item and ensuring that “*might*” poses a problem, account-giving takes place, which makes the problem clearer and signals that the problem has been recognized and accepted by the tester AY too. CI uses a different interactional resource in

providing the account as she translates the target sentence into Turkish by transforming the verbs in it into a form with “*might*” and pointing at each one of them with a pencil in her hand to emphasize the violated testing principle that “*might*” also fits well in the sentence. This makes the problem more explicit, leaving no room for further questioning and leads to its acceptance as well as the participants’ display of mutual understanding on the problem.

As for the institutional role distribution of the participants, there are indications of reviewer and tester roles and responsibilities in this IR interaction. Firstly, as the reviewer and problematizer of the item, CI orients to her right, responsibility and obligation to offer an account for the already presented problem while the tester AY is inquiring about the problem with her gaze on her exam paper (Heritage, 2013). In addition, since the problem is related to the context given in the text, which both participants have equal epistemic access to, the reviewer and the tester both carry out collaborative inquiry and thus, orient to not only the reviewer’s but also the tester’s responsibility and obligation to further evaluate the item and confirm the problem during the IR interaction (Heritage, 2013). Therefore, by acknowledging CI just as she initiates her turn of account-giving and employing choral co-production with CI, AY signals that she has fulfilled her responsibility to think about the matter at hand and has also recognized the problem. In this respect, this extract shows that the roles can be shared and the tester can act as a reviewer during the process of problematizing an item, taking time to review the item as a reviewer does and making an individual attempt (e.g., *bi bakıyım/let me see*) to try to understand the issue raised by the reviewer. In relation to this, the participants’ shared roles and active engagement in the task of problematizing the item is evident in their gaze behaviour as they both sustain their gazes on their exam papers throughout the interaction except when CI shifts her gaze to AY just prior to the completion of her account-giving in line 12. With their gazes on their papers, they implement some verbal and non-verbal actions such as reading the target sentence silently, translating it into Turkish, pointing at the verbs with a pencil, underlining “*should*” and “*might*” and producing a thinking face, which all signal the participants’ role as active agents during the phase of PI.

In extract 8a, the teacher AD teaching at B1 plus level reviews a multiple-choice reading comprehension item in a level achievement test of B1 plus groups and provides her feedback to the tester ME. The multiple-choice item asks which of the four options presents information that is true according to the text, which is about the health effects of moulds.

The reviewer AD problematizes the verb “*underrecognize*” given in the part of the text that the item is based on by stating that the students do not know the prefix “*under*” and thus, will not understand what “*underrecognize*” means, which will then prevent them from eliminating the options related to it and eventually finding out the correct answer. The tester ME also agrees with this.

#### Extract 8a. Moulds, Segment a

- 1 AD: >şimdi buna< hocalardan şöyle bi iti↑raz gelebilir  
*now an objection such as this might come to this from the teachers*  
 + both look at their exam papers  
 + points at a sentence in the text
- 2 (.) underrecognize'ı nerden bilcek (.)  
*how will they know "underrecognize"*  
 + ME shifts gaze to AD's exam paper  
 + turns gaze to ME
- 3 under prefix'iyle ilgili bi bilgi sahibi değil .hh  
*they have no knowledge about the prefix "under"*  
 + ME turns gaze to AD, then back to AD's paper
- 4 bur↑da (.) [bunu bilemezse ] (.)  
*here if they cannot know this*  
 + turns to the text and points at the sentence in the text
- 5 ME: [kelimeyi mi değiştirelim]  
*is it the word that we shall change*  
 + AD shifts gaze to ME, then back to her paper
- 6 AD: bunu bilemezse (.) şurda da er ona yakın ifade var  
*if they cannot know this, over there too there is an expression close to that*
- 7 (0.3)
- 8 <sup>1#</sup>health care professionals now ↑know that moulds can  
 + ME turns gaze to the text on a separate copy
- 9 cause allergies<sup>#1</sup> diye gidiyo (.)  
*and it continues*  
 + ME nods  
 #1 (lines 8-9) AD reads aloud the target sentence from the text and ME follows AD's reading on her copy of the text
- 10 [şimdi artık]  
*now*
- 11 ME: [now know ]
- 12 AD: biliyollar  
*they know*  
 + ME nods
- 13 (0.3)



- 14       bu tarz<sup>#2</sup> rahatsızlıkların sebebi (.) mould related  
15       olduğunu<sup>#2</sup>  
          *that the cause of these kinds of illnesses is mould-*  
          *related*  
#2 (lines 14-15) ME nods with her gaze sustained at the relevant part on her copy of the text  
16       (0.6)  
17       geçmişte o za↑ma:n (0.5) bilmiyollardı (0.2) gibi bi  
                                        + shifts gaze to ME  
18       [ifade çıkıyo             ]  
          *it means that in the past back then they didn't know*  
19   ME: [ya da şöyle dicesz]  
          *or we will say like this*

As the participants are both looking at their exam papers and the reviewer AD is pointing at a sentence in the reading text, extract 8a begins with AD's use of a preface, which creates a context for problematizing the item and projects the announcement of a problem by raising the possibility of teachers' objection to the item and presenting the problem on their behalf. This draws the tester ME's attention to the sentence pointed at by AD as she turns her gaze to AD's exam paper. Then AD implements the base FPP of problematizing the item in lines 2 and 3 by anticipating and voicing teachers' possible objection and pointing out what they would say with regard to the item. According to AD, the teachers might object that the students will not be able to understand the verb "*underrecognize*" due to their lack of knowledge about the meaning of the prefix "*under*". While articulating the problem, AD also puts an emphasis on the verb "*underrecognize*" and the prefix "*under*", further signalling that the problem is related to the use of this verb and its prefix in the target sentence in the reading text. Just as AD utters the verb "*underrecognize*", she shifts her gaze to ME, who then turns her gaze to AD, engaging in mutual gaze, and then quickly back to AD's exam paper through the end of AD's unit of talk.

The FPP of problematizing the item makes agreement or disagreement relevant next; however, without waiting for ME's response, AD shifts her gaze to the text, pointing at the sentence in question and extends her turn in line 4 with an emphasis and a rising intonation on the deictic "bur ↑ da / *here*" in turn-initial position. She considers how the students can deduce the meaning of "*underrecognize*" from the paragraph if they cannot know it. As a result, she brings a different dimension to the issue just raised. Just as AD initiates her turn, ME overlaps in line 5 and offers a candidate understanding (Antaki, 2012; Heritage, 1984b)

in the form of a confirmation request for the word that is to be changed, checking her understanding of the problem. Thus, she interprets the projected request for change in the item based on AD's statement of the problem and asks for its confirmation. This elicits AD's gaze but AD does not orient to ME's confirmation request and quickly turns back to her exam paper in an attempt to complete her ongoing turn about the students' way of deducing the meaning of "*underrecognize*" from the paragraph in the rest of the extract between lines 6 and 18.

In line 9, AD repeats the preliminary component of the conditional clause (*bunu bilemezse/if they cannot know this*) she produced in her prior not-yet-completed turn and then completes her ongoing turn by referring to a sentence in the reading text that could be a clue to understanding the verb "*underrecognize*" and reading it aloud with an emphasis on the time expression "now" and a rising intonation on the verb " $\uparrow$  know". At this point, ME turns her gaze to the text on a separate copy, follows AD as she reads aloud the sentence between lines 8 and 9 and shows agreement by nodding her head once AD finishes reading the sentence. Between lines 10 and 18, AD proceeds with her turn by making a connection between the sentence where the verb "*underrecognize*" is used and the following sentence that she has just read out. In doing this, AD partly translates the sentence she has read into Turkish in lines 10, 12 and 14, and explains the connection between the two antonyms "*know*" and "*underrecognize*" as well as the two different time periods expressed in the sentences, which would present a clue for the meaning of "*underrecognize*". AD states that professionals now know that the cause of illnesses such as allergies is mould-related as expressed in the sentence she has read aloud and this means that in the past they did not know this fact about moulds, leading to the meaning of "*underrecognize*". In line 17, AD also stretches the final sound of the time expression " $\text{za} \uparrow \text{ma} : \text{n} / \text{then}$ " and produces it with a rising intonation at the same time, emphasizing the time period conveyed in the sentence with "*underrecognize*" and then shifts her gaze to ME. As AD is making the connection between the two sentences, ME overlaps in 11, repeating the time expression and the verb (*now know*) in the sentence just read aloud and emphasized by AD, and with her gaze sustained at the relevant part on her copy of the text, she shows agreement by nodding her head between lines 12 and 15 as AD translates the sentence into Turkish. Finally, in line 19, without explicitly responding to AD's base FPP of problematizing the item, ME displays her acceptance of the problem by attempting to

change “*underrecognize*” given in the reading text through a projection of a suggestion for it (ya da şöyle dicesiz/or we will say like this), which makes the phase of SC relevant next.

Different from the other extracts, extract 8a shows how the way the problem is stated leads to its acceptance without being subject to further collaborative investigation and search for evidence by the group. The reviewer AD announces the problem on behalf of all the teachers of the group of students who will take the test by anticipating and voicing their possible objection to the item in terms of their students’ knowledge of the verb “*underrecognize*”. Thus, AD evaluates the item from the perspectives of both the teachers (iti↑raz gelebilir/an objection might come) and the students (nerden bilcek/how will they know, bi bilgi sahibi değil/they have no knowledge). In this way, she establishes common ground on which to base the problem and shares the responsibility for problematizing the item with the other teachers who are in equal epistemic position, presenting the problem as the product of shared understanding among the teachers. As a result, despite AD’s consideration of how the students can guess the meaning of the verb from the paragraph as an alternative, the tester ME directly accepts the problem and shows her readiness to change “*underrecognize*” by interpreting the projected request for change based on AD’s statement of the problem and offering a candidate understanding of it (Antaki, 2012; Heritage, 1984b) in line 5 as well as attempting to change the verb through a projection of a suggestion for it in line 19. In her confirmation request, she asks about what is to be changed rather than what causes the problem and goes beyond the phase of PI to that of SC, which displays her understanding and already acceptance of the problem. The tester might not want to take a risk as the problem is based on the teachers’ instructional events as well as their students’ second language knowledge, which is a domain of information that the tester has no direct epistemic access to. In addition, since the students’ lack of knowledge about the verb in question is explained with certainty in the form of a declarative statement (bi bilgi sahibi değil/they have no knowledge), the tester does not further question it and quickly shows her acceptance of the problem. This further shows that there are certain testing principles (e.g., test what you teach) that all the co-participants in the IR group adhere to in the construction of a language test and that they share a common ground in this respect.

With respect to the reviewer and tester roles and responsibilities in this extract, it is clear that AD orients to her role as the reviewer and teacher by presenting the violated testing principle from the teachers' and students' perspectives and evaluating the item considering the teaching context as well as the students' language knowledge based on what they are taught. AD, thus, indexes her epistemic right, responsibility and obligation (Heritage, 2013) to review the item with consideration of the context of learning and the needs of students and thus, to create the interplay between teaching and testing. Furthermore, AD also orients to her role as the problematizer of the item by bringing a different dimension to the raised issue and explaining how the students can deduce the meaning of the verb from the paragraph so that in the end the group can decide whether to keep the verb or to change it. Therefore, as the evaluator of the item, she signals her responsibility to consider all the dimensions of the problem and carefully think about it before deciding on it. On the other hand, ME's confirmation-seeking question for the projected request for change in the item and her quick attempt to change the verb indicate her focus on the elimination of the problem and the improvement of the item as the tester who is accountable for the test. In this respect, ME acts as a problem solver and approaches the item from the perspective of test construction (*değiştirelim/we shall change, şöyle dicese/we will say like this*). Additionally, while announcing the problem in line 2, AD shifts her gaze from her exam paper to ME, which orients to the tester's right, responsibility and obligation to know the problem as the writer of the item.

Finally, although this extract displays an orientation to the institutional role distribution of the participants, they happen to share the responsibility for their decisions as they are both engaged in the task of problematizing the item with their gazes sustained at the test materials (Goodwin, 1981; Rossano, 2013). ME, for instance, turns her gaze to AD's exam paper and follows AD as she reads aloud the sentence with her gaze on her copy of the text, nodding her head at certain points, which indicates that ME is also carefully thinking about the matter at hand and then showing agreement where necessary. Their equal engagement in the task is also evident in ME's use of an inclusive language when presenting the projected suggestion for change in the item during her turns-at-talk in lines 5 and 19 (*değiştirelim/we shall change, şöyle dicese/we will say like this*).

In extract 9a, the two teachers GE and DU teaching at B1 level review a multiple-choice reading comprehension item in a level achievement test of B1 groups and provide their feedback to the tester SA. The reading text that the multiple-choice item is based on is about young people becoming famous and how it affects their whole family. The multiple-choice item asks about the opinion of the famous striker Michael Owen's father with regard to his son's profession and his popularity. The stem states "*According to Michael Owen's father*" and the students are expected to complete this sentence with one of the four options given in the item. However, GE problematizes the distractor option "d" in this item, which states "*people can freely criticize Michael when he does not play*" as she believes that it could also be possible as a correct option when the students interpret the meaning of "can" as a possibility rather than a permission and that the item might end up having two possible correct answers. In the end, the other participants DU and SA also agree with the presented problem with regard to the distractor option "d".

#### Extract 9a. Freely Criticize, Segment a

- 1 GE: altı "c" mi "d" mi bilemedim  
*I'm not sure if six is "c" or "d"*  
 + all look at the item on their exam papers  
 + DU points at the item with a pencil in her hand
- 2 (0.3)
- 3 SA: ceyhanı hedeflemiştim [aslında (.) hocam ]  
*I actually aimed for "c" (coded with the name of a town/river in Turkey called "ceyhan") miss*  
 + turns gaze to DU's paper, then back to her paper
- 4 DU: [ben ceyhan demişim ]  
*I went for "c"*
- 5 (0.3)
- 6 dur bakıyım  
*let me see*  
 + bends forward to read option "d"
- 7 GE: <sup>1#</sup>people can freely criticize michael
- 8 when he doesn't play. <sup>#1</sup>  
 + shifts gaze to the text on the other page, then the item  
 #1 (lines 7-8) GE reads aloud option "d" from her exam paper and DU and SA follow her with their gazes on their papers
- 9 (1.0)
- 10 SA: <sup>2#</sup>aslında burda bad public reactions'dan <criticize'ı>  
 + points at the expression with a pencil
- 11 çıkarmasını: (.) bekledim  
*actually here I expected him/her to derive*

- "criticize" from "bad public reactions"**  
+ DU shifts gaze to the text on the other page, then back to the item
- 12 (0.5)
- 13 GE: [ay bu nerde geçiyo ya: ]  
**for god's sake where is this mentioned**
- 14 SA: [orda da (.) bunun (.) kötü bişey olduğunu] söylüyo  
15 ↑ya o yüzden (.) freely-<sup>#2</sup>  
**and you know it says there that this is something bad, that's why "freely"**  
#2 (lines 10-15) GE searches for the relevant sentence in the text by moving her finger on a paragraph and SA sustains gaze at the relevant paragraph in the text
- 16 DU: [beşinci paragrafta ]  
**in the fifth paragraph**  
+ shifts gaze to GE's paper
- 17 GE: <sup>3#</sup>[according to him the] ↑worst (.) p1- p1- [parts]  
+ SA shifts gaze to GE, then back to the sentence
- 18 SA: [huh ]
- 19 GE: of being famous are the BAD public reaction<sup>#3</sup>  
+ turns gaze to the item on the next page, then back to the text  
+ SA turns gaze to GE  
#3 (lines 17-19) GE reads aloud the relevant sentence in the text by moving her finger along it, and DU and SA follow her
- 20 (1.0)
- 21 yani hani (.) [people can]  
**well you know**
- 22 SA: [huh-huh ]
- 23 GE: freely (.) criticize michael  
+ shifts gaze to SA, then back to the sentence with her finger pointed at it
- 24 (0.6)
- 25 er <sup>4#</sup>which appear when the coach (.) sends (.) michael  
26 off the [field ]
- 27 DU: [ ° huh-huh ° ]<sup>#4</sup>  
#4 (lines 25-27) GE continues to read aloud the sentence by moving her finger along it, and DU and SA follow her with their gazes on the sentence in the text
- 28 (0.7)
- 29 GE: yani (.) hani (.) sahadan (.) michael'ı >attığında  
**well you know when he sends michael off the field**  
+ shifts gaze to SA + DU turns gaze to GE and nods
- 30 =yani> [michael OYNA]MADIĞIN[↑da  
**that is to say when michael does not play**  
+ shifts gaze back to the text
- 31 DU: [o kısım evet] [karşılıyo onu  
**that part yes corresponds to that**  
+ shifts gaze to SA, then back to her paper

32 (0.5)

33 GE: [er:

34 DU: [olabilir yani dediği doğru [GE'nin ]  
*it can be possible, I mean what GE says is true*  
+ SA shifts gaze to DU  
+ points at GE with a pencil in her hand

35 GE: [mı- mı-]

36 michael [o-

37 SA: [doğru (.) olur mu [yani?]  
*so would that be correct?*  
+ shifts gaze to DU's paper

38 GE: [oyna ] [madı]ğında (.)  
*when he does not play*

39 DU: [olur]  
*it would be so*

40 GE: kötü bir- (.) hani YU[:H yapıyolarmış ]  
*something bad- well they shout "boo"*  
+ shifts gaze to SA, then back to the text

41 SA: [aslında burda can'den]  
*actually here with "can"*  
+ turns gaze back to the item

42 benim (.) er amacım şeydi (.) hani babası (.)  
*my aim was the thing, well his father*  
+ circles "can" on her paper + shifts gaze to GE

43 ya (.) criticize etsinler bu iyi bişey (.) gibi bişey  
+ DU underlines the relevant sentence in the text  
+ shifts gaze to DU

44 demiş oluyo can diyerek  
*by saying "can" he happens to mean something like*  
*you see they shall criticize, this is something good*  
+ shifts gaze to GE

45 (0.4)

46 DU: öyle [di↑yo ]  
*he says so*  
+ turns gaze to the item on the other page

47 SA: [ama onu-] (.) onu [demi↑yo ]  
*but that- he does not say so*  
+ shifts gaze to DU, then to the text on her paper

48 DU: [°öyle di-° ]  
*so*

49 SA: aslında (.) [worst parts di↑yo ↑ya=  
*actually you know it says "worst parts"*

50 DU: [şi-

51 =<sup>5#</sup>the worst parts of being famous are the bad public  
+ GE produces thinking face and puts hand under her chin

52 reactions (.) which appear<sup>#5</sup>  
#5 (lines 51-52) DU reads aloud the sentence in the text and SA and GE follow her

53 (0.6)

54 GE: when the coach sends michael OFF [THE FIELD ]  
+ reads aloud the rest of the sentence and DU and SA follow her

55 DU: [off the field]  
+ nods head

56 buraya kadar (.) "d"nin (.) sanki olabilceğini  
+ underlines one part of the sentence

57 gösterir gibi. (.) ama  
*up to here it shows as though "d" can be like possible but*

58 (0.4)

59 and the lack of privacy (.) as he is recognized  
+ reads aloud the following sentence

60 wherever he goes (.) diyince  
*when it says*

61 "c" daha (0.2) şey (0.2) [basıyo gibi görünüyo]  
*it seems as though "c" is more like strong*  
+ shifts gaze to SA, then back to the item

62 GE: [tabi (.) "c" ama ] (.)  
*of course it is "c" but*

63 yani hani (.) oyna↑madığı zaman  
*well you know when he does not play*  
+ shifts gaze to SA

64 yani koç onu (.) oyun dışına attığı zaman  
*that is to say when the coach sends him out of play*  
+ shifts gaze back to her paper

65 (0.4)

66 insanlar (.) kötü- (.) yu::h diyebiliyolar=  
*people bad- can shout "boo"*  
+ DU turns gaze to the text on the other page  
+ shifts gaze to SA

67 =[yani freely ]  
*that is to say*

68 SA: [°huh° (.) anladım]  
*I understand*

69 GE: criticize ede[biliyolar gibi algılayabilir ]  
*s/he can perceive it as though they can criticize*  
+ SA shifts gaze to GE, then back to the item

70 SA: [işte anladım hocam ben orda can'ı]  
+ underlines "can" a couple of times



- 71 şey diye:- (.) aslında evet [can'ı koymamın hedefi- ]  
*you see I understand miss, there I put "can" so as*  
*to- actually yes, my aim for putting "can" is-*  
+ shifts gaze to GE
- 72 GE: [çocuk ama <onu ]
- 73 o:rdan> (.) çıkartamaz [gibi ]=  
*but it is like the kid cannot derive it from there*  
+ shifts gaze to SA
- 74 SA: [tamam]  
*okay*
- 75 GE: =yani ["d"ye de gider ]  
*so s/he can also go for "d"*
- 76 SA: [°değiştirelim°] evet hocam  
*we shall change it yes miss*  
+ turns gaze to her paper
- 77 =o zaman onu nasıl yapabiliriz ki  
*then how can we make it*

As all the participants are looking at the item on their exam papers, extract 9a begins with the reviewer GE's initiation of a pre-expansion sequence that projects a problem in the item and serves the base sequence of problematizing the item (Schegloff, 2007) by expressing her uncertainty about the correct answer and her indecisiveness between two options, that is options "c" and "d". This is met with a responsive uptake by the tester SA, who announces the expected answer "c" in line 3 by addressing GE in turn-final position with "hocam" and turning her gaze to DU's exam paper and then back to her own paper. In overlap with SA's turn, DU also responds to GE by articulating the option that she chose, which corresponds to the expected answer. After a gap of brief silence, DU bends forward to read option "d" as she asks to have a look at it in an attempt to check the likelihood of the other possible correct answer announced by GE. In line 7, GE takes over and reads aloud the problematic option, that is option "d" with a slightly falling intonation through the end of her reading as she shifts her gaze to the text on the other page in search of evidence and then back to the item. Meanwhile, the other co-participants follow GE with their gazes on their exam papers as she reads aloud the option.

Following a gap of silence, the tester SA takes a turn in line 10 and with her gaze sustained at the relevant paragraph in the reading text, she indicates her expectation with regard to the way the students should approach the option. In this way, she reflects on the item writing process, pointing out why she designed the option in such a way and how the option was

made relevant to the text. As SA indicates, in this option the students are expected to derive “criticize” from the expression “*bad public reactions*” given in the text, which she points at with a pencil. During SA’s turn-at-talk, GE searches for the sentence in the text which option “d” is based on by moving her finger on a paragraph and through the end of SA’s unit of talk, DU shifts her gaze to the text on the other page and then back to the item on her exam paper. Just as GE asks for the target sentence in the text with an expression of annoyance over her failure to find the sentence in line 13, SA extends her turn in lines 14 and 15 to explain how option “d” is then to be eliminated with her gaze still at the paragraph in the text. According to SA, criticizing is mentioned as one of the worst parts of being famous in the text and as evident in her use of “↑*ya/you know*” with a rising intonation, she displays this information as assumed and shared knowledge expressed in the reading text, which the co-participants all have equal access to. However, before she makes a connection between this information and the option, her explanation is cut off and she drops out when GE takes over to launch the base FPP of problematizing the item after finding the relevant sentence in the text. In line 16, DU also responds to GE’s request about where the target sentence is in the text by shifting her gaze to GE’s paper and articulating the paragraph number where the sought information is located.

In overlap with DU’s turn, having already found the sentence, GE implements the base FPP of problematizing the item between lines 17 and 40. In doing this, GE first reads aloud the first part of the sentence in lines 17 and 19 with a rising intonation on “↑*worst*” and an increase in the volume of “*BAD*”, which are key words for the option, by moving her finger along the sentence while the other co-participants follow her. At one point during GE’s reading aloud, SA shifts her gaze to GE and then quickly back to the sentence on her paper and shows listenership with “*huh*” in line 18. Just prior to the completion of GE’s reading in line 19, GE turns her gaze to the item on the next page and then back to the text while SA shifts her gaze to GE. After a gap of silence, before continuing to read the sentence, GE makes a connection between the part of the sentence she has just read aloud and the option in lines 21 and 23 by referring to the first half of option “d” (*people can freely criticize*), showing which part of the option the information in the sentence corresponds to. During GE’s turn-in-progress, SA shows listenership with the continuer “*huh-huh*” with her gaze at GE. Through the end of her unit of talk, GE shifts her gaze to SA and then back to the sentence with her finger pointed at it, moving on to read aloud the rest of the

sentence in lines 25 and 26 by moving her finger along the sentence as the other co-participants follow her with their gazes on the sentence and DU shows listenership with a continuer quietly uttered in line 27. Then shifting her gaze to SA, GE clarifies what is meant in the second half of the sentence she has just read aloud by translating it into Turkish in line 29 with frequent pauses (*sahadan michael'ı >attığında/when he sends michael off the field*) and then quickly translating the corresponding part in the option into Turkish in line 30 (*michael OYNAMADIĞIN↑da/when michael does not play*) with a shift of gaze to the text, making a connection between the sentence in the text and option “d”. While doing so, GE produces the verb “OYNAMADIĞIN↑da” given in the option at a higher volume and with a pitch rise on its final syllable, drawing others’ attention to it and showing that the option corresponds well with the information in the text.

During GE’s ongoing turn, DU turns her gaze to GE in line 29 and nods her head, displaying agreement with GE. DU also overlaps in line 31 and further acknowledges GE by (re)formulating GE’s account and coming to the conclusion that the meaning conveyed in option “d” matches with the sentence in the text, at which point she shifts her gaze to SA and then back to her paper. Following a gap of brief silence, just as SA shifts her gaze to DU, DU moves on to reconfirm GE’s claim with regard to option “d” in line 34 by (re)formulating the problem raised by GE and overtly acknowledging that option “d” can also be possible and thus, referring to the violated testing principle. In the same turn, she also explicitly supports and endorses GE’s remark with respect to the problem by confirming it as correct and ascribing the statement of the problem to GE (*dediği doğru GE'nin/what GE says is true*) while at the same time pointing at her with a pencil in her hand. In this way, DU affiliates with the stance adopted by GE in problematizing the item. In overlap with DU’s turn, GE also takes a turn at the same time with a turn-initial elongated hesitation marker and attempts to provide further clarification of the problem. In doing this, she repeats the Turkish version of the second half of option “d”, which she has just produced in her previous turn, and completes it with an interpretation of the “*bad public reactions*” mentioned in the text, that is the public shouting “boo” when Michael does not play. She, thus, provides a specific explanation for what is meant by “*bad public reactions*” in the text, at which point she turns her gaze to SA and then back to the text.

During GE's turn-in-progress, orienting to DU's acknowledgment of the problem with a shift of gaze to DU's exam paper, SA requests DU's confirmation with regard to the possibility of option "d" as the correct answer in the form of a candidate understanding (Antaki, 2012; Heritage, 1984b) of what DU has stated previously while displaying agreement with the problem. In this way, she seeks clarification and checks her understanding not only about the matter at hand but also about DU's stance toward the raised issue, eventually receiving a confirming response from DU in line 39 through the repetitional response "olur" (Heritage & Raymond, 2012). Then between lines 41 and 44, SA shows disalignment by prefacing her turn by "aslinda/*actually*", which projects some kind of disagreement, and provides further clarification of her aim in designing the option, which she could not fully explain previously when her utterance was cut off. In doing this, turning her gaze back to the item from DU's paper, she explains what is meant by the use of "*can*" in the option and thus, what actually the option states, displaying resistance to the problem announced by GE. According to her, with the use of "*can*", option "d" means that for the father, it is good if the public criticizes his son. While stating this, SA orients to the other co-participants by shifting her gaze to GE in lines 42 and 43, and to DU in line 43 as DU underlines the relevant sentence in the text. SA also circles "*can*" on her paper in line 42, signalling its role in the meaning of the option. SA's turn is met with a responsive uptake by DU, who presents a counter-argument by acknowledging that the meaning SA aimed to convey in the option is what is actually expressed in the reading text. DU increases her pitch through the end of her utterance and just prior to the completion of her turn, she shifts her gaze to the item on the other page. At this point, SA also turns her gaze to DU and then back to the text on her paper and shows disagreement and resistance by objecting to and challenging DU's assertion in line 47 with an opposing view that it is not the meaning expressed in the reading text.

Upon SA's disagreement, a non-minimal post-expansion sequence is employed where a process of elaboration and negotiation of meaning takes place between DU and SA with regard to what is expressed in the reading text as well as the option. In overlap with SA's turn, DU repeats her prior assertion in line 48 with a softer volume but before her turn is complete, she drops out as she is checking the matter at hand with her gaze at the item. Meanwhile, SA continues with her turn prefaced by "aslinda/*actually*" and provides an account in line 49 with a rising intonation by referring to the part of the sentence in the

reading text (i.e., *worst parts*) which presents contradictory evidence for the correctness of the option, meaning that criticizing is not favoured by the father. Evident in her use of “↑*ya/you know*” with a pitch rise, SA displays this information as assumed knowledge expressed in the reading text, which the co-participants all have equal access to. Without any time lapse and an orientation to SA, DU begins to read aloud the relevant sentence in the text in lines 51 and 52 while the other co-participants follow her and GE produces thinking face by putting her hand under her chin. When DU stops reading the sentence at one point, GE takes over and completes DU’s reading by reading aloud the rest of the sentence in line 54. Just as GE produces the final expression “*off the field*” in the sentence with a sharp increase in its volume, DU engages in choral co-production (Lerner, 2002) with the current speaker GE by co-producing this expression along with a head nod. In lines 56 and 57, DU then presents a counter-argument to SA’s account by underlining the sentence and pointing out that this part of the paragraph they have just read aloud seems to make option “d” possible, which provides evidence for her assertion about the option in question. She proceeds with her turn between lines 59 and 61 by reading aloud the following sentence about the “*lack of privacy*” and indicating that this part of the paragraph, on the other hand, makes “c” stronger as the correct answer. DU, thus, offers further evidence for both options and once again refers to the violated testing principle that there are two possible correct answers based on the information in the text.

Just prior to the completion of DU’s turn, she shifts her gaze to SA and then back to the item and GE overlaps to confirm DU that “c” is actually the correct answer with an emphasis on it. With a shift of gaze to SA, GE then moves on to explain once again that option “d” can also be possible by repeating her prior account between lines 63 and 69. In doing this, she first refers to the option, repeating the Turkish translation of the time clause in the option (i.e., *oyna↑madığı zaman/when he does not play*) with a rising intonation and then matching it with the corresponding part of the sentence in the text (i.e., *koç onu (.) oyun dışına attığı zaman/when the coach sends him out of play*) by shifting her gaze back to her paper. Then she repeats the outcome of this situation, that is Michael being sent off the field mentioned in the text in the same way as she previously explained by interpreting “*bad public reactions*” as people shouting “boo”. At this point, she shifts her gaze to SA while DU turns hers to the text on the other page. Following this, GE rushes to relate this information in the text to the main clause in the option in lines 67 and 69

with an emphasis on “freely” by explaining that the students can perceive “*bad public reactions*” given in the text as people freely criticizing Michael, which would then make option “d” correct. During GE’s turn-in-progress, SA shifts her gaze to GE in line 69 and then back to the item and displays and claims understanding in line 68 through the use of an acknowledgement token “°huh°” followed by an expression of understanding (i.e., *anladım/I understand*).

In overlap with GE’s turn just before it reaches completion, SA initiates a turn with an expression of understanding once again, which is; however, followed by a display of disagreement and resistance as she uses “*aslında/actually*” and starts to further explain her aim in the use of “*can*” in the option by underlining “*can*” a couple of times and shifting her gaze to GE but she cannot manage to complete her turn when GE overlaps and cuts off SA’s utterance with a shift of gaze to SA. Therefore, SA’s initial display of agreement is a “pro forma” agreement which serves to delay her upcoming dispreferred response as a contiguity-breaking practice and eventually, her use of an acknowledgement token as well as expressions or claims of understanding function as “pre-disagreements” (Schegloff, 2007). In lines 72 and 73, GE challenges SA’s remark and shows disagreement with her by providing an account from the students’ perspectives, stating that the students cannot derive the intended meaning from the option and thus, orienting to a different testing principle that is violated. This is received with SA’s agreement with the problem in line 74 through the use of a minimal expansion acceptance token as GE rushes to explicitly reassert the violated testing principle that option “d” can also be possible as correct. Upon this, SA displays understanding and acceptance of the problem as she agrees to change the option by turning her gaze to her paper and acknowledges GE through the compliance token “*evet/yes*”, eventually reaching a consensus on the issue. She then asks how they shall change the option, which projects the phase of SC as the relevant next action in the IR interaction.

Extract 9a demonstrates how the IR group resolves a dispute and reaches a consensus in problematizing a language test item through a process of account-giving and presentation of counter-arguments as a result of the emerging contradictory views and disagreement within the group. The discussion of the problematic option “d” starts with the tester SA’s pre-emptive explanation for the unlikelihood of the option in the pre-expansion sequence. It is at

this point where a connection is first made between “*criticize*” given in the option and “*bad public reactions*” in the text and the presentation of “*bad public reactions*” as “*the worst parts of being famous*” in the text is first considered as a clue for the father’s dislike of the situation, which eventually provides contradictory evidence for the distractor option “d” where criticizing is expressed with “*can*” as though it is favoured by the father. Thus, the tester pre-emptively explains her aim in the design of the option and considers the possibilities that might have been interpreted differently by GE in an attempt to prevent any misunderstanding before GE actually articulates the problem. In the base sequence, GE then announces the problem by making a connection between “*bad public reactions*” in the text and “*freely criticize*” in the option in the same way as SA has done and then between the time clauses “*when the coach sends Michael off the field*” in the text and “*when Michael does not play*” in the option. She also describes “*bad public reactions*” as people shouting “boo”, which corresponds to people criticizing in the option. Thus, she matches the option with the sentence in the text and regards the use of “*can*” as a possibility rather than permission as intended by the tester. As a result, the different interpretations of the option cause disagreement and dispute in the IR group.

The disagreement that arises leads to a process of negotiation of meaning between the tester and the reviewers in the rest of the interaction. SA, for instance, takes over to provide further clarification of her aim in the use of “*can*” in the option, explaining that different from what is expressed in the reading text, it makes the act of people criticizing his son acceptable for the father. This is, however, not received with agreement as the other reviewers insist on the problem and get engaged in a continuous cycle of providing counter-arguments to SA’s claim. DU and GE repetitively argue that the option corresponds well with the sentence in the text since criticizing is mentioned as one of the possible outcomes when Michael is sent off the field. On the other hand, SA objects to the reviewers and repeats her claim that what option “d” states is actually different from the sentence in the reading text since criticizing expressed through “*bad public reactions*” in the text is referred to as one of “*the worst parts of being famous*” and the use of “*can*” in the option conveys a meaning that is contrary to this, making the act of criticizing favourable by the father. Only when GE challenges SA’s argument from the students’ perspectives by stating that the students cannot derive the intended meaning from the option, which is a domain of information that the teachers have a greater epistemic access to than the tester, the discussion ends as SA understands the reviewers’ concern and accepts the problem without further challenging it. In other words,

GE manages to elicit a preferred response from SA only when she redesigns her prior account and changes its perspective (Pomerantz, 1984a; Pomerantz & Heritage, 2013) by orienting to a different violated testing principle, that is the expected way of eliminating the option being above the students' level and language skills.

As a result, as in extract 4a, this extract shows that upon SA's challenging of the problem, the participants collaboratively make an attempt to provide evidence and account and share their views until they reach a mutual understanding with respect to the problem. This also demonstrates that in such task-oriented workplace interactions, where the presented problem is multi-dimensional, contradictory view is not regarded as a threat to but as a source of validity and a way to resolve conflict (Davison, 2004). As Fulcher and Davidson (2007) indicate, during the validation process, there will be disagreements and alternative explanations that challenge the validity argument and test validation is about investigating those alternatives and successfully eliminating the counter-arguments to explain the evidence and support the use of a test for a given purpose (Fulcher & Davidson, 2007; So, 2014). Thus, such a collaborative way of understanding the problem and considering all the parameters of the problem in and through talk leads to its acceptance "in open discussion, through a process of dialogue and disagreement" (Fulcher & Davidson, 2007, p.21).

As in most of the extracts, extract 9a also reveals that problematizing a language test item is not simply about announcing the problem which is then directly accepted by the other co-participants as it involves some steps that the whole group collaboratively follows in and through interaction in the statement, explanation, and recognition of the problem. First of all, GE's initiation of a pre-expansion sequence serves the base sequence of problematizing the item by establishing the necessary pre-conditions and laying the groundwork (Schegloff, 2007; Stivers, 2013). In this pre-expansion sequence, GE learns what the expected answer is so that she can proceed and problematize the other possible correct answer. GE's pre-expansion sequence also leads to the tester's pre-emptive explanation of her aim just before the announcement of the problem and gives the tester the opportunity to clarify her aim in the design of the option and how it is expected to be eliminated. As a result, upon GE's expression of uncertainty about the correct answer, the tester's announcement of the expected answer along with an explanation for option "d" and DU's selection of this answer act as a go-ahead response (Schegloff, 2007) for problematizing the item and GE eventually launches the base FPP. Moreover, the pre-expansion sequence allows the participants to



orient to option “d” and get engaged in a preliminary inquiry (Pomerantz & Heritage, 2013) into the issue raised by GE as DU, for example, bends forward to read the option and asks to have a look at it.

Secondly, the base FPP of problematizing the item is realized through account-giving and presentation of evidence from the reading text. In doing this, GE makes a connection between the option and the sentence in the reading text by reading aloud the sentence in the text, translating both the sentence and the option into Turkish, describing what is meant by “*bad public reactions*” in the text and matching it with the option word by word. Therefore, GE employs some interactional features while articulating the problem to make her interpretation of the option clear and it is only after such a detailed explanation that the other co-participants respond to the base FPP of problematizing the item whereby DU displays agreement and SA shows disagreement with GE.

Thirdly, as the other co-participants listen to GE, they all get actively engaged in the task of problematizing the item as they constantly think about issue from different perspectives and follow a process of collaborative inquiry. This is most evident in the participants’ gaze behaviour since they continuously look at their exam papers and change their shifts between the reading text and the item on the other page only sometimes directing their gazes at each other for a short time, which signals that the participants actively think and inquire about the matter at hand in such a competing activity (Goodwin, 1981; Rossano, 2013). Even during their turns-at-talk, with their gazes on their papers, they also implement some non-verbal actions such as circling “*can*”, underlining the sentence in the text and producing a thinking face, which all indicate the participants’ role as active agents during the phase of PI.

As part of the collaborative inquiry, the participants evaluate option “d” from different dimensions, which leads to a careful thought on and analysis of the option with consideration of all the possibilities that could pose a threat to the validity of the item. While SA focuses on the use of “*can*” in the option and how it enables the option to deviate from the meaning conveyed in the text, GE and DU pay attention to the equivalence of the expressions in the option (i.e., *freely criticize, when Michael does not play*) with those in the reading text (*bad public reactions, when the coach sends Michael off the field*). GE also evaluates the option from the students’ perspectives and highlights the students’ incapability to derive the meaning intended by the tester from the option, bringing out a different testing

principle that is violated. At one point in the interaction, DU provides evidence not only for option “d” but also for option “c” by referring to the part of the sentence in the text that leads to the expected answer in order to show that there is information in the text that make both options possible. In addition, once DU agrees with the problem, DU and GE collaboratively offer an account in the post-expansion sequence in response to SA’s disagreement. They, for instance, do the reading of the target sentence through collaborative turn completions in lines 54-55 and employ choral co-production (Lerner, 2002) while DU is providing further account for the problem and as a follow-up, GE reproduces her prior account in support of DU’s explanation.

Finally, the other steps in problematizing the item involve SA’s confirmation request and the start of a non-minimal post-expansion sequence (Schegloff, 2007; Stivers, 2013) upon SA’s display of disagreement with the problem, which all play an important role in the elaboration, proof and recognition of the problem. Before explicitly expressing her stance about the issue, SA seeks clarification and checks her understanding of the other reviewer DU’s point about the matter at hand by requesting DU’s confirmation with regard to the possibility of option “d” in the form of a candidate understanding (Antaki, 2012; Heritage, 1984b) of DU’s prior utterance. Once SA realizes that they look at the issue from different perspectives and interpret the option differently upon DU’s confirming response, she starts to provide further clarification of her aim in the design of the option, which, in turn, paves the way for the rise of conflicting views, bringing to the surface each participant’s stance towards the issue. Similarly, the initiation of a non-minimal post-expansion sequence plays an important role in inviting further elaboration and account of the problem until a mutual understanding and agreement is reached (Stivers, 2013). The post-expansion sequence in this extract is, thus, characterized by the participants’ presentation of counter-arguments to each other’s claims and a process of negotiation of meaning with regard to what is expressed in the option, which eventually leads to the recognition and acceptance of the problem.

Regarding the institutional role distribution of the participants in this extract, there are indications of reviewer and tester roles and responsibilities in this IR interaction as in the case of other extracts. To begin with, SA orients to her role as the tester by going first in announcing the expected answer in line 3 and stating it from the perspective of test construction (*ceyhani hedeflemiştim aslında (.) hocam/I actually aimed for “c” miss*), and by reflecting on the item writing process and explaining her

aim in the way she has designed the item as well as her expectation from the students in lines 10, 11, 42, 70 and 71. This signals her right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item as the writer and thus, the 'owner' of the item. Thus, as the tester who has first-hand knowledge (Type 1 knowledge, Pomerantz, 1980) of the issue that has caused a problem in the item (i.e., the use of "*can*" in option "d"), she asserts epistemic priority over the matter by attempting to provide a clarification of her aim throughout the interaction. In addition, GE's and DU's shifts of gaze to the tester SA while announcing and explaining the problem and addressing her in particular indexes the tester's right to learn about the problem as the writer of the item so that the item she has prepared can be improved. On the other hand, the tester SA's shift of gaze mainly to the reviewer GE while providing her account in lines 42, 44, 71 and explicitly addressing her with "*hocam/miss*" in line 70 orients to GE's role as the problematizer of the item, who has the right to know the other participants' views in response to her statement of the problem, which makes agreement or disagreement relevant next.

With respect to the reviewer roles, in the pre-expansion sequence, before waiting for GE's announcement of the problem, DU makes an attempt to inquire about option "d" by bending forward to read it and asking to have a look at it. This indicates her role as the reviewer who shares the responsibility to identify a problem in the item with the other reviewer without leaving it to GE to first articulate it. In addition, DU's pointing at GE while displaying agreement with the problem and explicitly attributing the identification of the problem to her in line 34 ascribes an epistemic authority to GE and shows that DU might not want to assert epistemic priority, authority and control over an issue which GE has raised as the problematizer of the item (Heritage & Raymond, 2005; Stivers, 2005). However, at the same time, DU orients to her epistemic right to confirm GE in lines 31 and 34 by offering a (re)formulation (Heritage & Watson, 1979) of the problem and concluding from GE's remark that option "d" can also be possible as correct, which makes the problem more explicit. Therefore, she asserts epistemic authority and control, orienting to her role as the reviewer who is also responsible for evaluating and reviewing the item and who is, thus, in the same epistemic position as GE (Heritage & Raymond, 2005; Stivers, 2005).

Finally, in this extract, the reviewers and the tester together contribute to the process of problematizing the item by providing evidence from the reading text since they all have

equal epistemic access to the text and can judge the problem based on the available information and sources. This shows that the roles can be shared in such a task-oriented IR interaction depending on the availability of the accessible sources of information. On the other hand, GE frames her account not only based on the reading text but also from the students' perspectives in lines 69, 72, 73, 75 (*algılayabilir/s/he can perceive, çocuk ama >onu o:rdan< (.) çıkartamaz/but the kid cannot derive it from there, d'ye de gider/s/he can also go for "d")*) by anticipating how the students would be likely to approach the item as the teacher with access to the domain of teaching and the abilities and knowledge of her learners. In this way, GE orients to her role as the teacher and indexes her epistemic right, responsibility and obligation (Heritage, 2013) to offer an account with consideration of the context of learning and her students' skills and thus, to create the interplay between teaching and testing.

In extract 10a, the teacher YU teaching at B1 plus level reviews an item in a matching test task type designed for a story quiz of B1 plus groups and provides his feedback to the tester ME. The story exam tests students' comprehension of a story titled "Lorna Doone", which was assigned to them by the curriculum office to be read independently outside class. One of the tasks in the story exam is a matching test where a number of character names extracted from the story are given in a box with two extra names and just below the box of names, descriptions of four characters are listed. Students are to match each description with one of the character names given in the box by writing the name of the character in the blanks provided in front of the statements. During the reviewing of this matching test in and through interaction, the participants focus on the description of one particular character named Lizzie. The reviewer YU problematizes this item as Lizzie is one of the minor characters in the story whom he did not focus on while reading the story and thus, he was unable to match the statement with the expected answer "Lizzie". In the end, the tester ME also agrees with the presented problem.

#### Extract 10a. Lizzie, Segment a

- 1 ME: şunu zaten- (.) er birinciyi yapamamışsın.  
*that after all- you could not do the first one*  
 + points at the item with a pen  
 + YU looks at the matching task on his exam paper  
 + turns gaze to YU's paper
- 2 (0.3)

3 YU: tom (.) carver ve gwenny' ↑yi beğendim burda:  
***I liked tom, carver and gwenny here***  
+ points at the names one by one with a pencil  
+ ME shifts gaze to her paper, then nods and turns gaze back to YU's paper

4 1# (0.7)  
5 yani-  
***well***  
6 (1.2)  
7 hani ben (.) >açıkcası hiç<  
***you know frankly speaking I never***  
8 (1.0)  
9 kardeşlerinin karak (.) te-terlerine [çok tak] ılmadım  
***focused much on the characters of his siblings***  
+ shifts gaze to ME, then back to his paper

10 ME: [ °huh° ]  
+ smiles

11 (0.2)  
12 huh  
13 (2.0)  
14 er lizzie'den iki kere bahsediyö  
***it mentions lizzie twice***  
15 =onun cevabı lizzie  
***the answer to that is lizzie***  
16 (0.4)  
17 er[::]

18 YU: [lizzie tom'la şu (.) evlenen mi lizzie?  
***is lizzie the one who marries this tom?***  
+ points at "tom" in the box with a pencil  
+ shifts gaze to ME

19 ME: annie (.) tom' [la evlenen]  
***annie is the one who marries tom***  
+ YU turns gaze back to his paper

20 YU: 2# [hu:h ]  
21 (0.2)

22 ME: şöyle tom [ °faggus'la°  
***with the thing, with tom faggus***

23 YU: [>işte< çok (.) minor bişey ya ↑o (.)  
***you see that is something too minor***

24 ME: °hnm°  
25 (0.3)  
26 ya: şey diye anlatıyordu hat↑ta  
***well in fact it was explaining by saying the thing***  
27 (0.6)

28 hikayenin en sonunda bile işte lizzie de biriyle  
 29 evlenecek inşallah<sup>#1</sup> falan filan  
**even at the very end of the story, that lizzie too**  
**will marry somebody hopefully and so on**  
 + shifts gaze to YU, then back to YU's paper  
 #1 (lines 4-27) ME sustains gaze at the matching task on YU's exam paper  
 30 =[hani bu çok] (.)  
**you know this is too**  
 31 YU: [hu:h ]  
 32 ME: er asi ve kara- (.) şey u:m  
**rebellious and black- thing**  
 33 (0.5)  
 34 nası bi karakter diyim  
**how to say what she is like**  
 35 (0.4)  
 36 yaramaz bi karakter [o kız. ]  
**that girl is a naughty character**  
 37 YU: [anladım ]  
**I understand**  
 38 (0.8)  
 39 °unpredictable°  
 + ME produces thinking face and taps her fingers on the table  
 40 (0.8)  
 41 ME: ↑belki bire başka bişey  
**maybe for one something else**  
 + shifts gaze to her exam paper  
 42 (0.3)  
 43 YU: evet<sup>#2</sup>  
**yes**  
 #2 (lines 20-41) YU sustains gaze at the matching task on his exam paper

As both participants are looking at the matching task on their copy of the exam paper and the tester ME is pointing at the item with a pen, extract 10a begins with the tester's initiation of a pre-expansion sequence which projects a problem in the item, serving the base sequence of problematizing the item (Schegloff, 2007). In this pre-expansion sequence, finding out that the reviewer YU was unable to match the first description with any of the character names, ME anticipates a problem in the item and turning her gaze to YU's paper, pre-emptively announces YU's failure to answer the question before YU does so, seeking his confirmation. In response to this, in disalignment with ME's turn, YU refers to the three character names that he was able to match in line 3 by pointing at them one by one with a pencil and offers a positive assessment of them with the positive evaluation token "beğendim/*liked*". He

provides the sought-after confirmation in an indirect way by focusing on the character names he was able to match, leaving out the problematic name “Lizzie”, which is the expected answer for the first item, and thus, mitigating the projected problem through a positive assessment of those names first. During YU’s turn-at-talk, ME shifts her gaze to her paper and then nods her head, turning back to YU’s paper again. After a gap of silence, orienting to the first item this time, YU provides an account for his failure to match this item with any character name between lines 7 and 9 by reflecting on his own experience of reading the story and revealing that he did not focus much on the characters of the narrator’s siblings with an emphasis on the first person pronoun “ben/I”, which gives clues about the projected problem. Just as he articulates his account, he shifts his gaze to ME and then back to his exam paper and overlapping YU’s turn at a turn-final position, ME displays understanding through the quietly uttered acknowledgement token “huh” accompanied by her smiling in line 10 and another minimal agreement token in line 12.

Following a gap of silence, as she sustains her gaze at YU’s exam paper, ME takes over with a turn-initial hesitator and informs YU that the character Lizzie is mentioned twice in the story book in line 14. Then she rushes to interpolate a parenthetical insert into her turn in line 15 where she announces that the expected answer is “Lizzie”. By providing such additional information about the answer, she ensures shared understanding in the group and establishes common ground on which the problem can be based. Then just as ME continues with her turn with an elongated pause-filler, YU overlaps in line 18 and gets engaged in a preliminary inquiry (Pomerantz & Heritage, 2013) by addressing a confirmation-seeking question at the tester ME about Lizzie’s identity to check out if he can remember exactly who Lizzie is. In doing this, he recalls a marriage event from the story and pointing at the character “Tom” in the box with a pencil, he asks if Lizzie is the one who marries Tom with a shift of gaze to ME, soliciting ME’s confirmation. Then, ME provides a non-confirming response in line 19 as she elaborates through a non-conforming response (Raymond, 2003) that Annie is the one who marries Tom by putting an emphasis on Annie, at which point YU turns his gaze back to his paper. Overlapping ME’s turn at a turn-final position, YU displays understanding through an elongated acknowledgement token (hu:h) in line 20 while ME moves on to clarify the person whom Annie marries by giving his full name with softer volume in line 22. As a result, ME’s non-confirming response to YU’s preliminary inquiry acknowledges the projected problem, providing supportive evidence for it by indicating that the character Lizzie is a minor one and the item tests a detail that is difficult to remember by

the readers of the story. Thus, ME's disconfirmation encourages the production of the projected problem, serving as a go-ahead response for YU (Schegloff, 2007) and suggests an upcoming acceptance of the problem likely (Pomerantz & Heritage, 2013). Therefore, only after such a preliminary inquiry in the pre-expansion sequence, which proves YU's account in and through interaction, YU launches the base FPP of problematizing the item in line 23, orienting to the violated testing principle that the item tests a trivial content with an emphasis on the key word "minor" and a rising intonation on the recognitional demonstrative " $\uparrow$ o/*that*", which refers to Lizzie. He also presents his claim as assumed knowledge based on ME's non-confirming response through the use of "*ya/you see*". YU's announcement of the problem is received with ME's display of understanding via the quietly uttered acknowledgement token " $^{\circ}$ hmn $^{\circ}$ ", which is a "pro forma" agreement (Schegloff, 2007) that serves to delay her upcoming dispreferred response.

Following the base FPP of problematizing the item, which makes agreement or disagreement relevant next, ME makes her stance clearer about the matter at hand between lines 26 and 36 by indirectly showing partial resistance and disagreement with YU as both sustain their gazes at YU's exam paper. ME offers evidence and account with respect to what is mentioned about Lizzie in the story book, which acts as a counter-argument to YU's claim about the item. In a way, she provides a justification for her choice of the character Lizzie to be tested in the quiz by making reference to the shared product (i.e., the story book), which the participants both have equal epistemic access to and thus, to shared knowledge that is assumed to have been gained in a past event, that is the act of reading the story before the IR session (Can Daşkın & Hatipoğlu, 2019; You, 2015). In doing this, ME explains that even at the very end of the story, Lizzie's marriage with somebody was mentioned and puts an emphasis on "*en/the very*", signalling the likelihood of the information being remembered due to its presence at the end of the book. During her turn-in-progress, ME shifts her gaze to YU in line 29 and then quickly back to YU's exam paper. Then ME rushes through into another unit of talk where she describes Lizzie's character between lines 30 and 36 by first searching for the right adjective through frequent intra-turn pauses, the use of "*şey/thing*" and the hesitation markers "*er*" and elongated "*u:m*" as well as a self-initiation of repair in line 32, and then producing the repair solution in line 36 with a slightly falling intonation, coming up with the adjective "*yaramaz/naughty*" to describe Lizzie.



Just as ME starts to explain Lizzie's character, YU shows understanding and listenership through an elongated acknowledgement token "hu:h" in line 31 in overlap with ME's turn at a turn-initial position. Overlapping ME's turn at a turn-final position, YU then exhibits understanding more overtly with his gaze at his paper via an expression of understanding (anladım/*I understand*) and a (re)formulation (Heritage & Watson, 1979) produced with softer volume following a gap of silence where he offers an adjective (°unpredictable°) that would further express what ME means with respect to Lizzie's character. At this point, ME produces thinking face by tapping her fingers on the table. After the (0.8)-delay in responding with no uptake, without waiting for YU's further elaboration and account of the problem in response to ME's explanation, ME shows agreement with the problem and accepts it by suggesting to change the first item in the matching test with a turn-initial rising intonation in line 41, which projects the phase of SC as the relevant next action in the IR interaction. At the same time, she downgrades her suggestion with the hedge "↑belki/*maybe*" and shifts her gaze to her exam paper. This is met with a responsive uptake by YU, who exhibits agreement through the compliance token "evet/*yes*" in line 43.

Extract 10a demonstrates that problematizing a language test item in an IR interaction is not simply a matter of stating the problem in the item as it requires the use of some interactional features that lead to the collaborative identification, explanation and acceptance of the problem. In this particular extract, the implementation of a pre-expansion sequence where both participants are active agents especially plays an important role in the process of problematizing the item by establishing the necessary pre-conditions and laying the groundwork that serves the base sequence of problematizing the item (Stivers, 2013). First of all, different from the other extracts, the pre-expansion sequence in extract 10a is initiated by the tester, not the reviewer or the problematizer of the item as she pre-emptively projects a problem in the item before YU does so based on her observation that the first item in the matching test was left unanswered. This paves the way for a preliminary work towards the identification and recognition of the problem. The tester ME, for instance, announces the expected answer (i.e., Lizzie) in a parenthetical insert before YU asks for it and provides additional information about the character Lizzie by stating how many times she is mentioned in the story book, all of which present the necessary pre-conditions that might affect YU's evaluation of the item. In this way, ME also ensures shared understanding and

establishes common ground, based on which YU can proceed to the base sequence and problematize the item.

Secondly, in the pre-expansion sequence, the reviewer YU does not directly address the problematic item but first offers a positive assessment of the other three character names he could match in an attempt to mitigate the disaffiliative impact of his upcoming negative assessment with regard to the expected answer “Lizzie”. YU’s efforts to mitigate the projected problem continue as he gives clues about the problem by providing an account for his failure to match the first item based on his own personal way of reading the story book as evident in his use of an emphasized first person pronoun and revealing that he did not focus much on the characters of the narrator’s siblings. In this way, YU downgrades his upcoming remark while at the same time projecting the problem and laying the groundwork for its announcement because as the first assessor of the item, he might not want to index an implicit claim of epistemic priority (Heritage & Raymond, 2005). Thus, he uses a variety of resources to modify his epistemic stance in order to achieve epistemic congruence with the tester about an issue concerning the story book which they both have equal access to (Heritage & Raymond, 2005).

Finally, just before announcing the problem, YU gets engaged in a preliminary inquiry (Pomerantz & Heritage, 2013) as in extract 2a by asking ME if Lizzie is the one who marries Tom and requesting her confirmation, testing to see if he can recall the right information about Lizzie. However, ME’s non-confirming response indicates that Lizzie is a character not easily remembered and presents a minor detail from the story, which justifies YU’s failure to answer the question, proves his account that he did not focus much on those characters and leads to ME’s realization of the matter at hand in and through interaction. As a result, ME’s disconfirmation acts as a go-ahead response (Schegloff, 2007) for problematizing the item and they collaboratively contribute to the identification of the problem through such a preliminary inquiry in the pre-expansion sequence.

With respect to the way the base FPP of problematizing the item is implemented, YU announces his claim with greater certainty upon ME’s non-confirming response to his preliminary inquiry, which proves the character “Lizzie” to be a minor detail that is difficult to recall from the story. Therefore, throughout the interaction, YU’s account for the problem proceeds from specific to more general. While his preliminary account in the pre-expansion

sequence refers to a personal matter (i.e., kardeşlerinin karak(.)te-terlerine çok takılmadım/*I did not focus much on the characters of his siblings*), his account in the base FPP is framed in more general terms (i.e., >işte< çok (.) minor bişey ya ↑o/*you see that is something too minor*) after some preliminary work on the issue, orienting to the violated testing principle. In addition, YU presents his assertion in the base FPP as the violation of an assumed and shared testing principle that concerns the design of story exams in general within the institutional context in question (i.e., story exams should not test a minor detail from the story as students might have difficulty remembering those details). It is for the same reason why the tester ME does not object to the claim that a minor detail is tested by pointing out, for example, that such minor details should also be tested in story exams as much as the major ones. Instead, she objects to YU's remark that the tested character itself presents a minor detail by showing counter-evidence from the story with regard to Lizzie. This means that there are certain testing principles that all the co-participants in the IR group adhere to in the construction of a language test in their institution and that they share a common ground in this respect.

As for the way the base SPP is employed, ME exhibits a "pro forma" agreement (Schegloff, 2007) as she initially displays understanding and minimal agreement through the use of an acknowledgement token, which serves to delay her upcoming dispreferred response as a contiguity-breaking practice. However, after providing counter-evidence from the story with regard to the availability of sufficient information about Lizzie in the book and displaying partial resistance, she closes the sequence by this time showing agreement and directly accepting to change the item without waiting for any further elaboration of the problem by YU unlike in the case of other extracts characterized by emerging disagreement and dispute over the announced problem. This could be explained in a number of ways. First of all, since YU could not find the right answer as a teacher, it would be unfair to expect the students to answer the question correctly and thus, ME might not want to take a risk by disregarding the issue raised by YU even though she has an explanation for her choice of the character to be tested.

Secondly, ME might not have found the evidence she presented sufficient and strong enough to disprove YU's claim. This is most apparent in her difficulty in describing Lizzie's

character where she uses frequent intra-turn pauses, hesitation markers and a self-initiation of repair while searching for the right adjective to describe Lizzie. This may signal that she herself has difficulty recalling accurate information about Lizzie from the story and might have realized on spot that Lizzie is a minor detail. ME's slight falling intonation through the end of her turn of account in line 36 as well as her production of a thinking face by tapping her fingers on the table in line 39 might also be an indication of her display of uncertainty about testing the character Lizzie in the quiz.

Thirdly, YU's announcement of the problem prompts ME to justify her choice of such a character and elaborate on the item as the relevant next action. The reviewer YU has the right to know the tester's justification before moving onto the next phase of SC as it may change the course of action in a different way. However, the emergent delay with no uptake upon ME's counter-evidence alerts her to an upcoming dispreferred response and prompts her to back off and change her position by accepting to revise the item. This, in turn, results in YU's acknowledgement and affiliation with ME's stance, which is a preferred response that resolves the dispute and ensures contiguity in the IR interaction (Pomerantz, 1984a; Pomerantz & Heritage, 2013). In sum, ME's response to the base FPP shows that during the process of problematizing an item in and through talk, for whatever reason the participants are responsible to exchange any necessary information they possess about the item, which may or may not change the course of action or lead to new realizations online, and share their views in an open dialogue so that all the parameters of the problem can be considered and a mutual understanding can be reached in such a task-oriented interaction where the product is shared.

Regarding the institutional role distribution of the participants in extract 10a, there are indications of reviewer and tester roles and responsibilities in this IR interaction. To begin with, the tester especially makes her position and role clear throughout the interaction. For instance, on her own initiative, she announces the expected answer and gives additional information about it in the pre-expansion sequence, orienting to her epistemic right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item as the writer of the item. For the same reason, YU directs a confirmation request to the tester ME with respect to the information in the story book by shifting his gaze to her, which ascribes an epistemic authority to ME, orienting to her role as the tester with first-hand knowledge (Type 1 knowledge, Pomerantz, 1980) of the issue that has caused the problem in

the item (i.e., the choice of the character Lizzie). In fact, ME's immediate response without any delay, her reformulation of her response with the full name of one of the characters (line 22) as well as her non-conforming response (Raymond, 2003) that deviates from the terms and constraints of YU's polar question (line 19) prove her K+ status and epistemic authority (Heritage, 2013) as the tester, who has carried out a detailed analysis of the story book before preparing the quiz and thus, has a good command of the story. Therefore, ME also asserts epistemic priority over the matter by explaining on her own initiative how many times Lizzie is mentioned and what information is available about her in the book without being asked for it. Furthermore, the pre-expansion sequence is initiated by the tester herself upon her quick observation that the item was left unanswered, which orients to the tester's responsibility to be alert to any problems that might emerge in the IR interaction.

As for the reviewer roles in this extract, ME's announcement that YU could not do the first matching by using the past tense signals the reviewer's responsibility and obligation to have taken the exam as a student would before giving his feedback on the items (Heritage, 2013). In addition, YU's formulation of his confirmation-seeking question with some information he can recall from the story (i.e., marriage with Tom) as well as his offer of a (re)formulation with regard to ME's description about Lizzie's character demonstrate that YU has fulfilled his responsibility to read the story before coming to the IR meeting, orienting to his K+ status about the issue (Heritage, 2013).

In extract 11a, the two teachers BA and DE, who teach at B2 level, review a multiple-choice reading comprehension item in a level achievement test of B2 groups and provide their feedback to the tester EK. The reading text that the multiple-choice question is based on is about the running culture around the world, giving three examples of cultures which value running. The multiple-choice item asks the reason for why some monks in Japan run, which is one of the running cultures explained in the reading text, and lists in the options four reasons to choose from. In this item, the reviewer DE selects a wrong option, that is option "b", which states "*they want to achieve a higher status in their community*" while the other reviewer BA chooses the expected option "a", which states "*they want to appreciate the world better*". This raises dispute in the group with regard to the correct answer of the item as DE problematizes the expected answer by seeking evidence for it, whereas BA and the tester EK challenge the possibility of the option selected by DE and try to prove that it is option "a" and not "b" that is the correct answer. As a result, DE agrees with the other

participants and accepts that it was due to her carelessness and lack of attention to the part of the reading text that leads to the expected answer. Eventually, the item does not undergo any change and the IR interaction ends in this extract once the problem is resolved and eliminated.

### Extract 11a. Community

- 1 DE: dört "c:" beş "b:" (0.3) di mi?  
*four is "c" five is "b" isn't it?*  
 + EK and DE look at their exam papers  
 + BA turns over the page to look at the item  
 + shifts gaze to BA
- 2 EK: °huh-[huh°]  
 3 BA: [dört] "c".  
*four is "c"*  
 4 =beş (.) "B" Mî?  
*is five "b"?*  
 5 (0.2)  
 6 DE: evet.  
*yes*  
 + shifts gaze back to her paper  
 7 (0.3)  
 8 yanlış mı yapmışım  
*have I done it wrong?*  
 + turns gaze to EK  
 9 (1.0)  
 10 BA: they want <to appreciate> the world better (.)  
 + reads aloud option "a"  
 + DE turns gaze back to her paper  
 11 [demişim ben]=  
*I have said*  
 12 EK: ["↑a:" cevap]=  
*"a" is the answer*  
 + shifts gaze to DE  
 13 BA: ="a"  
 + shifts gaze to DE  
 14 (0.6)  
 15 DE: ben (.) niye ["b" yapmışım ]  
*why have I done it "b"*  
 + BA turns gaze back to the item  
 16 EK: [they want to achieve-]  
 + reads aloud option "b"  
 17 (1.0)

18 bi higher status peşinde-  
*in pursuit of a higher status*  
+ shifts gaze to DE, then back to her paper

19 BA: evet  
**yes**

20 (0.3)

21 ben "a" demişim.  
*I have said "a"*

22 DE: niye world better'a git-  
*why go for "world better"*

23 ((EK shifts gaze to DE, then back to her paper and  
turns over the page to look at the text while BA and  
DE look at the item on their papers))

24 higher status orda ben mesela onu şöyle infer ettim  
*I, for example, inferred it that way there*  
+ turns over the page to look at the text  
+ EK shifts gaze to DE, then back to the text  
+ BA shifts gaze to DE and turns over the page to look  
at the paragraph

25 (0.4)

26 spiritual insight'a higher status- (.) ruhlarında  
+ points at "spiritual" in the text and shifts gaze to EK

27 higher status olarak algıladım  
*higher status to spiritual insight- I perceived it  
as higher status in their spirits*  
+ BA and EK both shift gaze to the item on the next page

28 (2.0)

29 BA: in their [community diyo ama ]=  
**but it says**  
+ DE turns gaze to BA, then back to the text on her exam paper  
+ turns gaze back to the text and then to DE

30 EK: [↑community var a↑ma:]=  
**but it says "community"**  
+ shifts gaze to DE

31 DE: =nerde.  
**where**  
+ EK shifts gaze back to the item

32 (0.4)

33 EK: cevap[ta  
**in the answer**  
+ turns gaze to DE

34 BA: [seçenekte cevapta  
**in the option in the answer**  
+ EK shifts gaze back to the item

35 (1.0)

36 DE: ↑huh spiri:tually  
+ EK shifts gaze to DE, then to her paper

37 (0.5)

38 öyle mi?  
*is that so?*  
+ turns over the page to look at the item  
+ BA shifts gaze to the paragraph

39 (0.2)

40 BA: tabi  
*of course*

41 (3.0)

42 DE: huh community'yi bilmezse nolcak  
*what if s/he does not know "community", what will happen*  
+ EK shifts gaze to DE, then to her paper  
+ shifts gaze to EK

43 (0.5)

44 in their socie-

45 BA: ama community'yi bil[sin artık ya: ]  
*but well s/he should know "community"*  
+ DE turns gaze to BA  
+ BA and EK both shift gaze to DE, then back to their papers

46 DE: [bilir (.) di mi?]  
*s/he would know it, wouldn't s/he?*

47 (0.7)

48 bi dakika ben "a"ya hiçbi şekilde niye gitmedim?  
*one minute why have I not by any means gone for "a"?*  
+ shifts gaze back to the item and marks option "a" on her paper  
+ BA shifts gaze to DE's exam paper, then back to hers

49 (0.5)

50 er yani nerden-  
*I mean from where-*

51 (0.5)

52 they want to appreciate life better'a nerden götürdü  
*where did you get to*  
+ reads aloud option "a"  
+ DE and EK both turn over the page to look at the text

53 =yani [hangi cüm]le götürdü sizi  
*I mean which sentence took you to*  
+ shifts gaze to EK

54 BA: [ ° huh ° ]

55 ((all look at the text on their papers in search of the sentence key to option "a"))



56 EK: <purs:(.)ue-> (.)<sup>1#</sup>[er (.) (monks) will develop a  
57 greater admiration and sensitivity to life. ]  
58 BA: [işte (.) the monks will develop  
**here it is**  
59 a greater (.) admiration and sensitivity to life.]<sup>#1</sup>  
+ shifts gaze to DE,  
then back to the paragraph  
#1 (lines 56-59) EK and BA simultaneously read aloud a sentence in the text while DE scans  
through the text on her paper  
60 (0.4)  
61 DE: nerde ya? bi dakika=  
**oh where is it? one minute**  
62 EK: =yedinci paragraf  
**seventh paragraph**  
63 ((DE searches for the sentence by moving a pen along  
the paragraph while BA and EK sustain their gazes at  
the paragraph))  
64 DE: huh ben tabi [oraya bakma]mışım ki (.) bi dakika  
**but of course I have not looked there, one minute**  
65 BA: [ortalarında]  
**in the middle of it**  
+ EK shifts gaze to DE, BA and back to the text respectively and then to  
DE and back to the text again  
66 ((DE continues her search for the sentence by moving  
a pen along the paragraph))  
67 DE: °nerde°  
**where is it**  
68 ((finds the sentence and underlines its first part))  
69 huh  
70 (0.3)  
71 the monks will develop a greater admiration and  
+ reads aloud the sentence by moving her pen along it  
72 <sensi↓tivity>  
73 (1.0)  
74 [to life ]  
+ DE and EK both turn over the page to look at the item  
75 BA: [or- (.) ordan] gidiyoruz yani.  
**so we go from there**  
76 ((DE looks at the item on her paper while EK changes  
gaze between the paragraph and the item twice and BA  
does something on her phone))  
77 DE: °appreciate the world°  
+ turns over the page to look at the target paragraph  
78 ((EK shifts gaze to the item and then back to the  
paragraph))

79 °admiration and sensitivity to life°  
+ reads from the paragraph in the text and then turns over the page to look at the item  
+ EK turns gaze to the item

80 ((DE and EK look at the item on their papers and then as DE sustains gaze at the item, EK shifts her gaze to the paragraph on the other page while BA continues to do something on her phone))

81 [bilmiyorum ben-]  
**I do not know**

82 EK: [unravel the ] secrets of the universe deeper  
+ reads aloud from the paragraph  
83 meaning of life  
+ shifts gaze to the item  
+ DE shifts gaze to EK

84 ((DE turns over the page to look at the paragraph while EK shifts gaze to DE's paper and then back to the item on her paper and BA picks her exam paper to look at the paragraph))

85 DE: huh  
86 (0.4)  
87 bütün bir paragraf olarak mı.  
**is it the paragraph as a whole**  
+ EK shifts gaze to DE, then to her paper and back to DE again

88 (0.4)

89 BA: e tabi they also say that they ↑now can  
**yes of course**  
+ reads aloud from the paragraph in the text  
+ EK shifts gaze to BA, then to her paper

90 un[ravel the secrets of-]

91 DE: [huh see hear ] taste smell  
+ reads aloud from the paragraph in the text  
+ EK shifts gaze to DE, then to her paper

92 (0.7)

93 tamam ya: doğru doğru  
**oh okay right right**  
+ turns over the page to look at the item and circles option "a"  
+ BA turns over the page to look at the item

94 (0.3)

95 ben demek ki o paragrafa hiç (0.2) bakmamış mıyım?  
**this means have I not looked at that paragraph at all?**  
+ BA and EK shift gaze to DE, then EK turns gaze back to her paper

96 =doğru şimdi zaten bi tek cümle değil  
**right, now after all it is not one single sentence**  
+ turns over the page to look at the text again  
+ BA shifts gaze to the item



rushes to acknowledge EK's response without any time lapse by repeating the expected option "a" at a higher volume in line 13 and thus, affiliating with EK's stance.

Once the other participants meet on a common ground upon the tester's articulation of the expected answer, DE questions why she has gone for option "b" in line 15 with her eye gaze on the item. Then EK overlaps DE's turn at a turn-final position just as BA directs her gaze towards the item and starts to read aloud option "b" in order to figure out why DE might have selected this option. As a result, following a gap of brief silence, EK makes an attempt to provide an account for the unlikelihood of option "b" by referring to "higher status", which is the focus of the option, and indicating that in the text the runners in question are not after a higher status with a shift of gaze to DE. However, before she completes her account in line 18, she drops out as she directs her gaze back to her paper, further checking the matter at hand. Despite EK's incomplete account, BA displays agreement with EK by confirming her account via the compliance token "evet/yes" in line 19 followed by her restatement of the option that she chose, which further signals her affiliation with the tester while at the same time conveying an oppositional stance towards DE.

Following the reviewers' provision of two different answers, the tester's announcement of the expected answer that corresponds to one of them and her account for the incorrectness of option "b", this time DE orients to option "a" in line 22 and challenges it by asking which information in the text leads to "world better" given in the option and thus, seeking evidence for the correctness of option "a". However, before she completes her turn, she drops out as she is trying to understand option "a" with her gaze sustained at the item. A gap of silence ensues with no uptake as EK shifts her gaze to DE and then back to her paper, turning over the page to look at the text in search of the sought-after evidence for option "a" while DE and BA are looking at the item on their exam papers. On the other hand, between lines 24 and 27, turning over the page to look at the reading text, DE provides an account and justification for her choice of option "b" by offering evidence from the reading text and explaining her interpretation of the option in relation to the text with a shift of gaze to EK in line 26. According to her, "higher status" in the option corresponds to "spiritual insight" given in the reading text and thus, the latter has prompted her to perceive the former in terms of spirituality. During her turn-in-progress, EK and BA both shift their gazes

to DE, showing listenership, and then back to the relevant paragraph in the reading text and DE points at “*spiritual*” in the text, which is the key word for option “b” according to her.

Upon DE’s account for option “b”, just prior to the completion of her turn, EK and BA both shift their gazes to the item on the next page, getting engaged in an inquiry into the issue raised by DE with regard to option “b”. As a result, following a gap of silence, EK and BA both challenge DE’s argument in lines 29 and 30 by presenting counter-evidence from the option, that is the availability of the expression “*in their community*” in the option, which enables it to deviate from the meaning expressed in the reading text. In doing this, directing their gazes towards DE, EK and BA employ choral co-production (Lerner, 2002) as they co-produce “*community*” at the same time in the same manner and end their turn with “*ama/but*”, which further marks their turn as a counter-argument to DE’s account for option “b”. Based on this, BA puts an emphasis on the word “*community*” whereas EK increases her pitch in word-initial position, drawing DE’s attention to it. In response to this, since DE is looking at the text but not the item, she cannot find where “*community*” is mentioned and quickly asks BA and EK for its place in the test material, at which point EK shifts her gaze back to the item and then turning her gaze to DE, responds to her in line 33 by stating that it is present in the answer. In overlap with EK’s turn, BA also responds by first articulating that it is in the option and then repeating EK’s response turn as EK shifts her gaze from DE to the item. After a gap of brief silence, this is met with a responsive uptake by DE, who first shows agreement through an acknowledgement token uttered with a rising intonation in word-initial position and then repeats the adverb form of the key word in the text for option “b” (*spiri:tually*), which elicits EK’s gaze to DE. This is followed by DE’s use of a newsmarker (*öyle mi?/is that so?*) in the form of a confirmation request in line 38 just as she turns over the page to look at the item where “*community*” is used. At this point, BA shifts her gaze to the paragraph and provides confirmation through the use of the marked confirmation “*tabi/of course*” (Stivers, 2011) in line 40, displaying certainty with regard to their prior claim about option “b”.

Following a gap of silence while the participants are all looking at their exam papers, just as EK shifts her gaze to DE and then back to her paper, DE takes over to counter-challenge EK and BA’s argument against option “b” in line 42 from the students’ perspective by playing the devil’s advocate, presenting the possibility of students’ lack of knowledge with respect

to the word “community” and asking about its outcome. During her turn-at-talk, she shifts her gaze to EK and then attempts to offer “society” as a suggestion to replace “community” in line 44 in case students do not know it. However, her utterance is cut off when BA responds to DE’s counter-challenge in line 45 by strongly claiming that “community” is a word the students should already know as evident in the emphasis placed on the verb “*bilsin/s/he should know*” and the stretching of the final word in her turn. At this point, the participants engage in mutual gaze as DE shifts her gaze to BA while BA and EK direct their gazes towards DE and then back to their papers. Overlapping BA’s turn at a turn-final position, DE immediately backs down, acknowledging BA’s remark with regard to the students’ knowledge of “community” in line 46 through a partial modified repeat (Kim, 2002; Stivers, 2005) of BA’s prior turn (*bilir/s/he would know*) followed by a confirmation check. In this way, they reach a consensus on the incorrectness of option “b” and on its appropriacy as a distractor which the students can eliminate.

Once the participants achieve mutual understanding on option “b”, shifting her gaze back to the item and marking option “a” on her paper, DE again orients to option “a” in line 48 by questioning why she has not by any means gone for it. Just prior to the completion of her turn, BA directs her gaze towards DE’s exam paper and then back to hers. When a gap of brief silence ensues, DE proceeds with her turn to rephrase her question twice between lines 50 and 53, seeking evidence for option “a”. In doing this, in line 52, she launches a self-initiated self-repair, reformatting her turn in line 50 by reading aloud option “a” and asking more specifically what information in the text led EK and BA to choose this option. During her turn-in-progress, she and EK both turn over the page to look at the text in search of evidence as prompted by DE’s question. In line 53, too, DE rephrases her question once again to make it even more specific by soliciting the particular sentence in the text that led to option “a” with a shift of gaze to EK. In overlap with DE’s turn, BA shows understanding of her question through an acknowledgement token (*huh*) that is uttered with a softer tone of voice.

As all the participants are looking at the text on their papers in search of the sought-after sentence that is key to option “a”, EK takes over in line 56 in order to read aloud the target sentence in the text. In doing this, she first reads the word “*purs : (.) ue*” slowly with an elongation and a micro-pause in the middle of its articulation, which displays EK’s

hesitation with regard to its status as the sought-after evidence in the text but then she cuts off her reading aloud and following a hesitation marker, having found the exact sentence she has been looking for, she uses a self-initiated self-repair and reads aloud another sentence in the same paragraph in lines 56 and 57. BA engages in choral co-production (Lerner, 2002) with the current speaker EK by co-producing the target sentence and through the end of her choral co-production, she shifts her gaze to DE and then back to the relevant paragraph in the text. Meanwhile, DE scans through the text on her paper in search of the sentence that has just been read simultaneously by EK and BA but unable to find where it is, she asks for its place in line 61 and EK immediately responds to her by articulating the number of the paragraph where the sentence is available (*yedinci paragraf/seventh paragraph*). Then, DE starts to search for the sentence in the seventh paragraph by moving a pen along it while EK and BA sustain their gazes at the paragraph. As she orients to the seventh paragraph, she realizes that she has not looked at this paragraph in line 64. While DE continues her search for the sentence between lines 64 and 67, BA further guides her by indicating that the sentence is in the middle of the paragraph in line 65, at which point EK orients to the other co-participants by directing her gaze towards them and then back to the text. Being unable to find the sentence, DE quietly asks for it once again in line 67 as she is still scanning through the paragraph and eventually finds the sentence in line 68, underlining its first part and acknowledging her identification of the sentence with “huh”.

Between lines 71 and 73, DE reads aloud the sentence by moving her pen along it and through the end of her reading aloud, she produces the word “<sensi↓tivity>” slowly with a sharp falling intonation and completes her reading after a gap of silence. Just prior to the completion of her turn, DE and EK both turn over the page to look at the item in an attempt to make a connection between the sentence DE has just read aloud and option “a”. Overlapping DE’s turn at a turn-final position, BA makes it clear that this sentence is the one that leads to option “a” in line 75 with a slight falling intonation. DE continues to look at the item on her paper while EK changes her gaze between the paragraph and the item twice and BA does something on her phone. As DE is thinking about the matter at hand, she reads the key part of the option with a softer volume (i.e., °*appreciate the world*°) and then turns over the page to look at the target paragraph and read the main part of the sentence from the paragraph quietly (i.e., °*admiration and sensitivity to life*°) in line 79, and looks back at the item again, trying to make a connection between the evidence

in the text and the option. Meanwhile, EK changes her gaze between the item and the paragraph repeatedly. Following this, as DE sustains her gaze at the item, EK shifts her gaze to the paragraph on the other page while BA continues to do something on her phone.

In line 81, DE takes over and provides a hedged response with the disclaimer “bilmiyorum/*I do not know*” to EK and BA’s remark with regard to option “a”, which still exhibits some kind of disagreement and disaffiliation with their stance. However, DE’s utterance is cut off when EK overlaps her turn at a turn-initial position and offers a second piece of evidence by reading aloud another sentence from the same paragraph in the text in lines 82 and 83 and then shifting her gaze to the item. At this point, DE directs her gaze towards EK and then turns over the page to look at the paragraph as prompted by EK’s new evidence while EK shifts her gaze to DE’s paper and then back to the item on her paper and BA picks her exam paper to look at the paragraph. Upon DE’s thought on the second sentence just read aloud by EK, she comes to the realization that the paragraph as a whole, in fact, presents clue and evidence for option “a” and thus, checks her understanding through a candidate understanding (Antaki, 2012; Heritage, 1984b) framed in the form of a confirmation request in line 87 with an emphasis on “paragraph/paragraph” following her use of the acknowledgement token “huh” in line 85, which elicits EK’s gaze to DE. This is received with BA’s marked confirmation via “*tabi/of course*” (Stivers, 2011) and EK’s shift of gaze to BA and then back to her paper. In lines 89 and 90, BA continues her turn by reading aloud the sentence, a part of which EK has just read aloud, but her turn is cut off when DE overlaps in line 91 and with a turn-initial acknowledgement token, reads aloud a part of another sentence from the same paragraph, which presents further evidence for option “a”. At this point, EK shifts her gaze to DE and then back to her paper.

After a gap of silence, just as DE turns over the page to look back at the item and circles option “a”, she displays a stance of acceptance of this option and stronger agreement via the acceptance token “*tamam/okay*” followed by the elongated change-of-state token “*ya:/oh*” (Heritage, 1984b) and her confirmation of the other co-participants’ claim with regard to option “a” through the stance marker “*doğru doğru/right right*” in line 93. Through the end of DE’s unit of talk, BA also turns over the page to look at the item. In line 95, DE proceeds with her turn and states her realization that the paragraph that is key to



option “a” has escaped her notice while answering the question, which elicits the gaze of both EK and BA to DE. Then just as she turns over the page to look at the text again, DE rushes to once again confirm the other co-participants’ assertion as correct via the stance marker “doğru/*right*” and display her uptake of it in lines 96 and 97 by offering a (re)formulation of the account the others provided for the expected correct answer and stating the conclusion that she has drawn from their discussion. She explains that not one single sentence but the whole paragraph presents information that corresponds to option “a”. Then between lines 98 and 100, scanning through the paragraph by moving her pen along it, she quickly repeats the specific parts of the paragraph that are key to option “a” (i.e., *smell, un(.)ravel the secrets of the universe*) and shows her acceptance of this option as the correct answer through her use of the turn-final as well as the turn-initial minimal expansion acceptance token “tamam/*okay*” in lines 100 and 102 respectively. This is followed by her articulation of the elongated change-of-state token “*ya:/oh*” and her reconfirmation of the other co-participants’ claim with regard to option “a” through the stance marker “doğru doğru/*right right*” in line 102. As a result, the dispute that initially arose in the IR interaction is resolved in and through talk and the seemingly problematic item ends up being accepted as it is when the group agrees that option “a” is, in fact, the correct answer, posing no problem.

Different from the other extracts analysed so far, extract 11a demonstrates how an initially seemingly problematic language test item is decided to be kept as it is without undergoing any change as a result of resolving an emerging dispute in the IR interaction with regard to the item. Thus, in this particular extract, the item turns out to have no problem and the discussion revolves around not proving a problem in the item as in the other extracts but arguing for the validity of the item through a continuous process of account-giving and evidence-seeking. It is clear that the problem is not pre-identified by the reviewers while taking the test and the item is at first assumed to have no problem since with her initial confirmation request for the answer of the item, DE’s aim is not to problematize but to approve the item. However, the problem emerges within the IR interaction in and through talk in the pre-expansion sequence when the reviewers announce their answers for the item and reveal that two different options were selected, leading to the rise of a dispute within the group with regard to the correct answer of the item. It turns out that DE has selected a wrong option (i.e., option “b”) while BA has chosen the expected answer (i.e., option “a”),

affiliating with the tester's stance. Hence, DE's initiation of a pre-expansion sequence before approving the item in the base sequence enables her to check the necessary pre-conditions (Stivers, 2013). The recipients' non-confirming responses lead to the emergence of a dispute over the correct answer, acting as a blocking response for approving the item (Schegloff, 2007). As a result, DE problematizes the expected answer by seeking evidence for it and has second thoughts about the two options, whereas BA and the tester EK disagree that there is a problem in the item and challenge the possibility of the option selected by DE, trying to prove that it is option "a" and not "b" that is the correct answer.

The participants follow a particular interactional pattern while discussing the two options as they first reach an agreement on the incorrectness of the distractor option "b" where they agree that the expression "*in their community*" makes the option unlikely and then on the appropriacy of option "a" as the correct answer which is located in one of the paragraphs in the reading text that went unnoticed by DE. The interactional process for the discussion of option "b" involves EK's initial account-giving for the unlikelihood of option "b" as the correct answer, DE's explanation of her interpretation of option "b" in relation to the text, BA and EK's collaborative challenging of DE's explanation by providing counter-evidence through choral co-production, DE's playing the devil's advocate and making a counter-challenge towards BA and EK's argument from the students' perspective, and BA's formulation of another counter-challenge to DE's questioning, which eventually prompts DE to back down and agree on the incorrectness of option "b". Thus, the discussion on option "b" undergoes a continuous process of providing account and evidence and challenging each other.

On the other hand, the interactional process for proving option "a" as the correct answer involves DE's seeking evidence for it, BA and EK's provision of evidence collaboratively from the reading text through choral co-production, DE's display of partial resistance, and then EK's presentation of a second piece of evidence from the same paragraph, which leads to DE's realization that one of the paragraphs that is key to option "a" has escaped her notice and her understanding that the paragraph as a whole provides clues for the correct answer. Only upon such a realization, DE shows agreement and displays a stance of acceptance of option "a" as the correct answer. Thus, the participants' discussion on option "a" paves the way for a collaborative identification and recognition of evidence for this option, which eventually resolves the dispute that has arisen with regard to the correct answer of the item.

In sum, this extract shows that such a collaborative way of understanding the matter at hand and considering all the parameters of the two options in and through talk “in open discussion, through a process of dialogue and disagreement” (Fulcher & Davidson, 2007, p.21) leads to the identification of the source of the problem (i.e., DE’s selection of an unexpected option) that is DE’s misinterpretation of option “b” and her lack of attention to one of the paragraphs in the text and to the realization that there is no threat to the validity of the item. In this way, their open exchange of ideas and information enables the co-participants to reach a mutual understanding about the item. As Fulcher and Davidson (2007) indicate, during the validation process, there will be disagreements and alternative explanations that challenge the validity argument and test validation is about investigating those alternatives and successfully eliminating the counter-arguments to explain the evidence and support the use of a test for a given purpose (Fulcher & Davidson, 2007; So, 2014).

Regarding the institutional role distribution of the participants in this extract, there are indications of reviewer and tester roles and responsibilities in this IR interaction. To begin with, DE’s shift of gaze to the tester while seeking confirmation for her answer in line 8, providing account for her choice of option “b” in line 26, challenging the other co-participants’ argument in line 42, seeking evidence for option “a” in line 53 as well as the tester’s provision of the expected answer in line 13, and her quick account for the incorrectness of option “b” in line 18 all orient to the tester’s epistemic right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item as the writer of the item.

As for the reviewer roles, DE’s challenging of BA and EK’s counter-argument in terms of the students’ knowledge of the word “*community*” in the option (i.e., *community’ yi bilmezse nolcak/what if s/he does not know “community”, what will happen*) signals her role as the teacher who has the responsibility to review the test items with consideration of the context of learning and her students’ skills. Based on this, BA goes first in responding to DE’s challenging by reflecting on her students’ knowledge and claiming that they would know the word, which further orients to her epistemic status and shared role as the teacher with access to the domain of teaching and the abilities and knowledge of their learners. DE’s modified repeat of BA’s assertion with regard to the students’ knowledge at the same time indexes her epistemic right to confirm BA, asserting

epistemic authority and control as the teacher who is in the same epistemic position as BA (Heritage & Raymond, 2005; Stivers, 2005). In addition, by using “*tabi/of course*” in lines 40 and 89, BA treats the presented information as assumed knowledge, implying that the state of affairs indexed by DE’s confirmation request is self-evident and BA’s position on the matter is previously held (Heritage, 2002; Stivers, 2011). In this way, she asserts epistemic priority and orients to her role as the reviewer who has answered the question correctly in line with the tester’s expectation. Furthermore, the reviewers’ reflection on their test-taking process by using the past or present perfect tense (i.e., *yapmışım/I have done*,  *demişim/I have said*, *infer ettim/I inferred*, *algıladım/I perceived*, *gitmedim/I have not gone*, *götürdü/took you*, *bakmamışım/I have not looked*) signals their responsibility and obligation to have taken the exam as a student would before giving their feedback on the item (Heritage, 2013).

Finally, despite the institutional role distribution, the participants happen to share their roles and responsibilities in the IR interaction as they are all actively engaged in the task of resolving a dispute about an issue which they all have equal epistemic access to (i.e., the reading text as the source of information). This is most evident in their gaze behaviour as they constantly turn over the page and change their gazes between the reading text on one page and the item on the other page, actively thinking and inquiring about the matter at hand (Goodwin, 1981; Rossano, 2013). BA and EK’s choral co-production (Lerner, 2002) is also an indication of their collaborative engagement in the task. BA does not just leave it to EK to argue for the validity of the item as the tester who is accountable for the test but collaboratively contributes to the process of resolving the emerging dispute with regard to the answer of the item.

In extract 12a, the two teachers YE and CI teaching at B1 level review an item in a vocabulary quiz and provide their feedback to the tester SA. They focus on an item in a vocabulary passage presented in the form of a binary-choice task where one alternative word is given for each target word that is tested. The students are expected to choose the word that is appropriate for the context given in the text. The reviewers collaboratively problematize one of the words, that is the adjective “*ambitious*” tested in the vocabulary passage from various perspectives and in terms of three violated testing principles. They claim that (1) “*ambitious*” did not appear in the textbook and is not specifically studied in class and thus,

the students may not know it; (2) the sentence where “*ambitious*” is used is difficult to understand, and (3) two points are tested in one single sentence. They all agree with the identified problems in the item and in the end, decide to omit the item instead of revising it as it is found too problematic. Thus, the reviewing interaction for this item ends in this extract once it is decided to be left out.

#### Extract 12a. Ambitious

- 1 YE: <onikiyi> biraz- ambitious: (.) er: nerde geçiyodu:  
*twelve is a bit- where did "ambitious" appear?*  
 + all look at the item on their exam papers  
 + produces thinking face and puts hand on her forehead  
 + CI shifts gaze to YE
- 2 (0.9)  
 3 ya da geçecek miydi  
*or was it going to appear?*  
 4 =ambitious öğretecek miydik (.) [hatırlayamadım]  
*were we going to teach "ambitious"? I could not remember*  
 + shifts gaze to CI  
 + SA shifts gaze to YE
- 5 CI: [ambitious ] (.)  
 6 personality adjective (0.2) de  
*is in "personality adjective"*  
 + SA shifts gaze to CI
- 7 =>ama nerde geçiyodu<  
*but where did it appear?*
- 8 (0.3)  
 9 YE: sanki [ben onu- ]  
*I think I it-*
- 10 CI: [bi sonraki] kitapta [mıydı  
*was it in the next book?*
- 11 SA: [personality daha evet  
 + looks at the book open in front of her  
 + YE shifts gaze to SA's book
- 12 burda yok sanki=  
*yes it seems like "personality" is not here yet*  
 + shakes head
- 13 CI: =yok [(°evet°) ]  
*not here yes*  
 + turns gaze back to her paper

- 14 SA: [bilmezler] mi  
**wouldn't they know it?**  
+ shifts gaze to YE
- 15 YE: [evet ]  
**yes**  
+ turns gaze back to her paper
- 16 (0.3)
- 17 yani onu bilmeyebilir[ler=  
**well they might not know it**  
+ shifts gaze to SA, then back to her paper
- 18 SA: [hu:h  
+ turns gaze back to her paper
- 19 YE: =sonra öğretmez galiba ambitious'ı  
**I think we will teach "ambitious" later**
- 20 (0.8)
- 21 SA: [um: ne diyebilir]riz  
**what can we say?**
- 22 YE: [but if you are-]  
+ reads from the target sentence
- 23 (2.3)
- 24 CI: bana da bu biraz (0.3) anlaması zor geldi (.)  
**I also found this a bit difficult to understand**
- 25 yani ambitious'ın ötesinde de aslında  
**I mean actually beyond "ambitious"**  
+ YE nods head
- 26 (1.7)
- 27 YE: çünkü bi de (.) şey gibi sanki (.) aynı cümleden iki  
+ SA and CI turn gazes to YE, then to their papers  
+ points at the two items in the  
same sentence respectively with a pencil
- 28 soru çıkarmış çı gibiyiz (0.2) [SA ]  
**because it is also like the thing, like as if we have  
derived two questions from the same sentence**  
+ shifts gaze to SA, then back to her paper
- 29 CI: [evet]  
**yes**
- 30 YE: bak [hani (.) but (.) şey- ]  
**look you know like**  
+ turns gaze to SA's paper and points at "but" with a pencil
- 31 SA: [en azından but'la bağlan]mış  
**at least it is connected with "but"**  
+ points at "but" on her paper with a pencil
- 32 =hani bunun nokta koyup however'la da bağlayabiliriz  
**well we can put comma here and connect it with  
"however"**  
+ YE shifts gaze to her paper

33 a↑ma (.) [eğer sıkıntılı] var diyosanız  
**but if you think there is a problem**  
 34 YE: [evet ]  
**yes**  
 35 SA: bunu da atabiliriz.=  
**we can omit this**  
 36 CI: =bunu atalım bence  
**I think we should omit this**  
 + shifts gaze to SA  
 37 (0.3)  
 38 SA: tamam  
**okay**  
 39 ((all cross out the item on their papers))

As all the participants are looking at the item on their exam papers, extract 12a begins with YE's initiation of a pre-expansion sequence where she gets engaged in a preliminary inquiry (Pomerantz & Heritage, 2013) by questioning hesitantly where the adjective “*ambitious*” appeared with reference to past instructional events accompanied by her display of a thinking face with her hand on her forehead. Through the end of her unit of talk, the other reviewer CI shifts her gaze to YE. After a gap of silence with no uptake, YE continues her preliminary inquiry by rephrasing her question twice in the form of a self-initiated self-repair in lines 3 and 4 and then claiming being unable to remember whether “*ambitious*” was or would be taught. She shifts her gaze to CI just as she rushes to reformulate her question the second time in line 4, maintaining mutual gaze with CI who is already looking at her and then the tester SA also directs her gaze towards YE. YE redesigns her initial question by changing its point of reference from shared past (*nerde geçiyodu:/where did it appear*) to future instructional events (*geçecek miydi/was it going to appear, öğretecek miydik/were we going to teach*), asking whether “*ambitious*” would appear and whether they would somehow teach it in the future. In her second rephrasing of her question in line 4, she also uses an inclusive language with respect to the teaching of “*ambitious*” in class, prompting the other reviewer to consider their teaching plans, which she is expected to have a direct access to and which are taken as the base in syllabus-based language test construction.

Overlapping YE's turn at a turn-final position, CI takes over in line 5 and responds to YE's question by informing that “*ambitious*” takes place in personality adjectives, at which point SA shifts her gaze to CI. Without any time lapse, CI then quickly displays uncertainty about

where personality adjectives were studied by asking in a fast pace where the word appeared in line 7 and then guessing if it is in the next book they are to study in line 10. Thus, although CI knows the topic area “*ambitious*” falls into (i.e., personality adjectives), YE cannot get an exact response from CI, who is also uncertain about where personality adjectives or “*ambitious*” in particular will be taught, further inquiring about it as YE does. During CI’s ongoing unit of talk, YE overlaps in line 9 and attempts to state her guess with regard to the place of “*ambitious*” in the program but she drops out as CI proceeds with her turn. Just prior to the completion of CI’s turn, looking at the textbook which is open on the table in front of her, SA provides a negative response to YE’s preliminary inquiry which serves as a go-ahead response (Schegloff, 2007) for problematizing the item. SA asserts that personality adjectives are not yet within the scope of the book accompanied by her head shake. At the same time, she mitigates and downgrades her claim with the hedge “*sanki/I think*”. Just as SA orients to the textbook, YE also shifts her gaze to SA’s book during her unit of talk. Upon the tester’s assertion, CI quickly confirms SA in line 13 through a partial modified repeat (Kim, 2002; Stivers, 2005) of SA’s turn where she repeats the negative response token “*yok/not*” with a shift of gaze back to her paper and then YE also acknowledges SA through the compliance token “*evet*” in line 15 by turning her gaze back to her paper.

In overlap with the other co-participants’ turn, directing her gaze towards YE, SA seeks confirmation from her with regard to the students’ knowledge of the adjective in the form of a negative interrogative tilted toward a no response (*bilmezler mi/wouldn’t they know it*). After a gap of brief silence, YE then provides a confirming repetitional response (Heritage & Raymond, 2012) to SA’s confirmation request in line 17 with a shift of gaze to SA first and then back to her paper while at the same time launching the base FPP of announcing the problem that the students may not know the adjective and rushing to explain that it will be taught later in the program in line 19. YE also downgrades her claim for the problem through the mitigating qualification of possibility expressed in “*bilmebilirler/they might not know it*” and “*galiba/I think*”. Overlapping YE’s turn, SA shows agreement in the base SPP via the elongated acknowledgement token “*hu:h*” by shifting her gaze back to her paper. Following a gap of silence, with a turn-initial elongated hesitation marker, SA seeks for a suggestion in line 21 to change the item using an inclusive language, which further signals her acceptance of the



problem articulated by YE and projects the phase of SC as the relevant next action in the IR interaction. Simultaneously with SA's turn, YE reads aloud one part of the sentence where "*ambitious*" is used but she drops out without completing her reading and a gap of silence ensues with no uptake.

In line 24, CI takes over to problematize the item further from a different perspective by stating that she found the sentence where "*ambitious*" is used difficult to understand and that the problem is not just about students' knowledge of the word. YE shows agreement with CI's announcement of a different problem by nodding her head during CI's turn-at-talk. After a gap of silence, YE takes a turn in line 27, carrying out a third round of problematizing the item as she presents another violated testing principle that two points are tested in one single sentence by pointing at the two items in the same sentence respectively with a pencil. At this point, SA and CI both shift their gazes to YE and then back to their papers. While articulating the third problem in the item, she explicitly addresses the tester SA with her name in turn-final position and directs her gaze to her. Once again she downgrades her assertion by using hedges (*sanki/as if, gibi/like*) and an inclusive language (*gibiyiz/as if we*), seeming as if they together constructed the item in such a way.

Overlapping YE's turn at a turn-final position, CI quickly shows agreement through the compliance token "*evet/yes*". Then YE attempts to provide further clarification in line 30 by orienting to SA's paper and pointing at "*but*" in the target sentence with a pencil to make her point clearer but her utterance is cut off when SA overlaps in line 31 and displays resistance and disagreement. SA argues against YE's remark that the sentence is at least connected with "*but*" with her pencil pointed at "*but*" and then rushes to offer a suggestion that the sentence can be broken into two with "*however*" as a way to easily overcome the problem newly presented by YE. Thus, SA does not regard this problem as serious and YE agrees with her suggestion in line 34 (*evet/yes*) with her gaze at her paper while SA's turn is already in progress. However, in lines 33 and 35, launching a *but*-prefaced turn uttered with a pitch rise, SA moves on to suggest omitting the item this time as a result of a number of other emerging problems in the item. This is met with an immediate responsive uptake by CI, who accepts SA's suggestion through a full modified repeat (Kim, 2002; Stivers, 2005) of SA's turn by shifting her gaze to SA. SA then uses the acceptance token

“tamam/okay” in line 38 and they all together cross out the item on their papers, which indicates their agreement with and acceptance of the problems discussed with regard to the item. As a result, the reviewing of this particular item ends in this extract without going through the other phases in the IR interaction.

Different from the other extracts, extract 12a demonstrates how a language test item is problematized collaboratively from a number of perspectives and then decided to be omitted from the test. The problematizing process is carried out in three rounds through the use of different interactional resources. The first one involves a collaborative investigation as prompted by YE’s initiation of a preliminary inquiry (Pomerantz & Heritage, 2013) in a pre-expansion sequence as in extracts 2a and 10a, which serves the base sequence of problematizing the item by establishing the necessary pre-conditions and laying the groundwork (Schegloff, 2007; Stivers, 2013). Although the problematizer YE had enough time to think about the item and her teaching experience before the IR interaction, she does not, for instance, directly state her assertion regarding the problem but instead gets engaged in a preliminary inquiry (Pomerantz & Heritage, 2013). In this way, she seeks concrete evidence for the projected problem by directing a question to the group with reference to the shared past instructional events (Can Daşkın & Hatipoğlu, 2019; You, 2015) where the adjective “*ambitious*” tested in the vocabulary passage appeared as well as its future teaching. This evidence is acknowledged with SA’s negative response to the preliminary inquiry, which encourages the production of the projected problem, acting as a go-ahead response (Schegloff, 2007) and suggests an upcoming acceptance of the problem likely (Pomerantz & Heritage, 2013). Thus, only after such a preliminary inquiry in the pre-sequence, YE implements the actual act of problematizing, which is then welcomed by the other co-participants. In a way, through this question, YE invites a collaborative identification of the problem and manages to involve all the co-participants in problematizing the item, which eventually results in the direct acceptance of the problem in the group. Throughout the process of the preliminary inquiry, the other co-participants CI and SA also pose some other questions that pave the way for further careful thought on the item as they seek evidence for the problem. For instance, CI makes guesses about the place of “*ambitious*” in the program by questioning whether it appears in the next book they are to study (*bi sonraki kitapta mıydı/was it in the next book*) and SA solicits confirmation for students’ present knowledge state with respect to the adjective (*bilmezler mi/wouldn’t they know it*). Thus, in and through the talk-in-

interaction, they collaboratively show an orientation to the testing principle that was violated (i.e., that the item tests what was not taught) and YE ensures shared understanding and common ground on which to base the problem.

The second round of problematizing the item is initiated by the other reviewer CI, who announces an additional problem that she pre-identified while taking the test (*anlaması zor geldi/I found this difficult to understand*). With her statement of the problem, the problem proceeds from the tested word itself to the sentence where the word is positioned, posing a greater threat to the validity of the item. This is because whatever the tested word is, if the sentence is confusing and difficult to understand, the item will not reach its aim in testing the target word, not supporting the use of the test for a given purpose. Finally, the third round of problematizing the item is launched by YE again, who orients to a different violated testing principle by drawing the others' attention to the layout of the item. CI's focus on the sentence might have enabled YE to realize that the same sentence also tests two words, which is against the testing principles. Therefore, the talk-in-interaction may lead to new realizations with respect to the item, affecting the final decision in the IR interaction. However, different from the first two rounds of problematizing the item, YE's announcement of the problem is met with the tester's resistance and disagreement as she suggests that the sentence can be broken into two and the words can still be tested as such. Nevertheless, her disagreement with the last stated problem does not prevent her from disregarding the other emerging problems as she eventually makes the final decision by offering to omit the item. Thus, not the problematizers but the tester herself makes such an offer based on her evaluation of the presented problems, which poses a threat the validity of the item to a great extent (i.e., the tested word is not studied in class and the sentence is difficult to understand) and which does not seem to leave much room for change and improvement in the item. As a result, this extract shows that each participant has a different contribution to make to the process of problematizing the item as they identify and discuss various problems by considering all the parameters of the item, help each other to recognize and prove a problem, and present their views with regard to the acceptability of the problem, which all play an important role in the final decision about the item.

As for the institutional role distribution of the participants in extract 12a, there are indications of reviewer and tester roles and responsibilities in this IR interaction. To begin with, in the preliminary inquiry, YE shifts her gaze to CI while reframing her question from

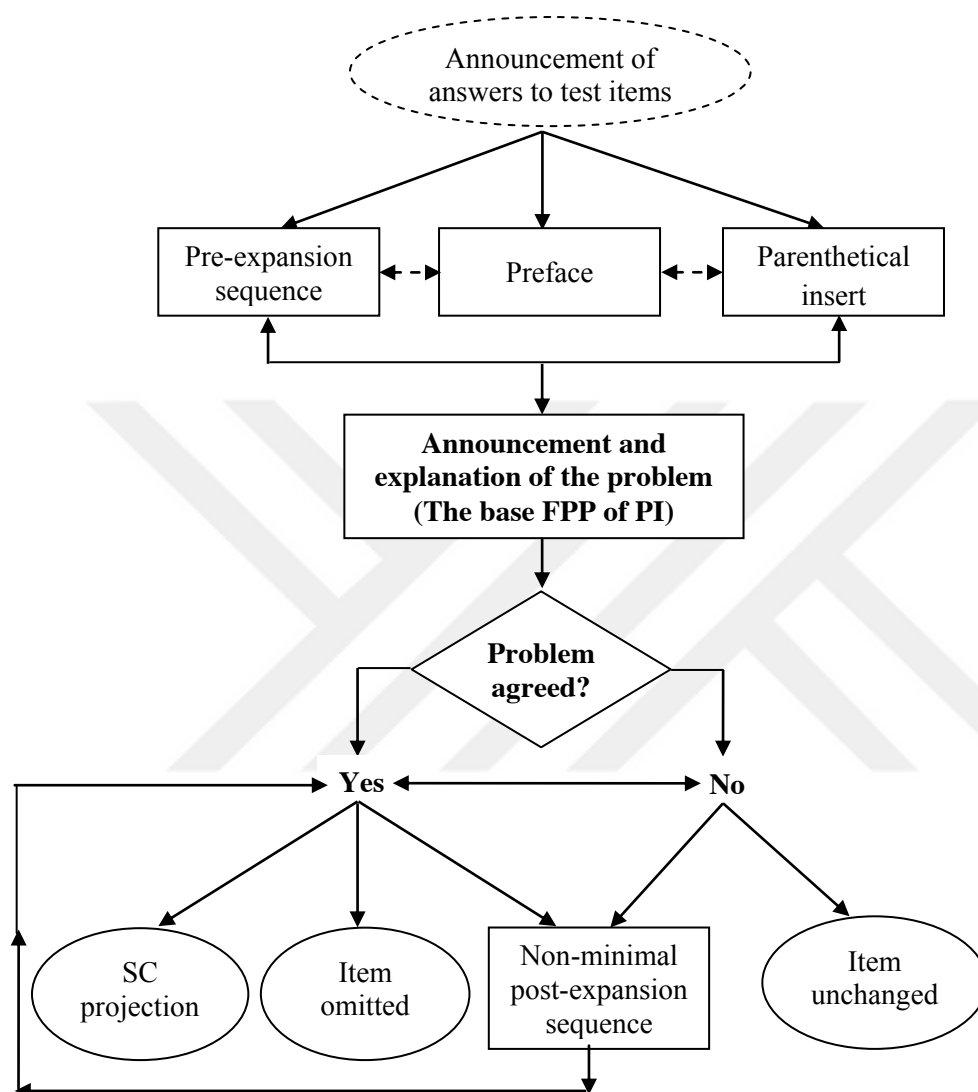
the teaching perspective because as the reviewers and teachers teaching at the same proficiency level following the same program, CI is in the same epistemic position as YE and they share the same institutional responsibilities. Similarly, SA's request for confirmation from YE with regard to the students' knowledge of the target word and YE's repetitional response orients to YE's role as the teacher who has access to the domain of teaching and the abilities and knowledge of her learners. On the other hand, while problematizing the item the second time in line 28, YE explicitly addresses SA this time with her name in turn-final position and shifts her gaze to SA, orienting to SA's role as the tester who has prepared the test and her right to learn about the problem so that they can work towards resolving the problem in the item. In addition, CI's use of a partial modified repeat of SA's assertion in line 13 rather than simply confirming it through an agreement token signals her epistemic right to confirm the tester and to announce the problem as the reviewer of the test, indexing epistemic authority over the matter at hand (Heritage & Raymond, 2005; Stivers, 2005).

Although this extract displays an orientation to the institutional role distribution of the participants, they happen to share the responsibility for problematizing the item and for their final decision in such a task-oriented IR interaction where the product is shared. The co-participants' equal engagement in the task of problematizing the item is most evident in their gaze behaviour as they continuously look at their exam papers only sometimes directing their gazes at each other for a short time, actively thinking about the matter at hand no matter what their role is in the group (Goodwin, 1981; Rossano, 2013). In addition, the co-participants are often seen to use an inclusive language in the IR interaction, which orients to their shared roles and shared epistemic knowledge. For instance, in line 4, YE asks CI whether they will teach "*ambitious*" using the inclusive "*we*" pronoun (*öğretcek miydik/were we going to teach*), which indicates YE and CI's shared roles as teachers with equal epistemic access to the domain of teaching. Likewise, SA's use of an inclusive language while seeking for a suggestion to revise the item in line 21 (*ne diyebilirsiniz/what can we say*) and offering a change to overcome one of the problems identified in the item in line 32 (*bağlayabiliriz/we can connect it*) demonstrates that all the co-participants have the epistemic right, responsibility and obligation (Heritage, 2013) to contribute to the item change process and that they jointly decide on the changes to be made in the item. YE uses an inclusive language in line 28 too

while problematizing the item the second time with reference to an item writing practice and a testing principle (aynı cümleden iki soru çıkarmış gibiyiz/*like as if we have derived two questions from the same sentence*) where she seems as if she was also involved in the test construction process and prepared the item in such a way together with the tester. This not only displays an orientation to the teacher's understanding of the item writing practice within the institutional context in question and to her approach to the item as a tester with consideration of the testing principles for item writing but also depersonalizes the presented problem, opening up a space for a collaborative understanding of the problem (Waring, 2017).

What is more, even though YE has the right to problematize the item as the reviewer of the test, she often uses mitigating qualifications and hedges throughout the IR interaction in order to downgrade her claims and mitigate the disaffiliative impact of the problem which directly concerns the tester's practice in designing the test (i.e., selection of a word not studied in class, testing two words in the same sentence). In this way, as the first assessor of the item, YE might not want to be too assertive with her negative assessments about the item and to implicate epistemic priority about a matter that is directly related to the tester's practice (Heritage & Raymond, 2005) but rather invite careful thought on the problem by the other co-participants, who are all to act as reviewers and problematizers in and through talk. For instance, YE uses the hedges "*gibi/like*" and "*sanki/as if*" in line 27, which express less certainty with regard to her remark about the item although it is based on something observable since it is clear that two points are tested in one single sentence. As a result, despite their distinct institutional roles, the participants use a variety of resources (e.g., inclusive language, hedges) to modify their epistemic stances, signalling their shared roles as teachers, reviewers, testers and problematizers of the test items.

#### 4.1.1. Summary of Interactional Patterns in Phase PI



\*See Appendix C for the meaning of flowchart symbols

**Figure 3. Overall Structural Organization of Phase PI**

Regarding the findings for the overall structural organization of the phase of PI in the language test IR interactions analysed for this study, as Figure 3 illustrates, it is seen that the base FPP of PI is not directly instantiated but it is preceded by one or a combination of the following interactional resources: (1) the reviewers' announcement of their answers to the language test items as the sequence-initiating action (extracts 1a, 3a, 4a, 6a, 7a, 9a, 11a),

which might be met with the other co-participants' display of surprise at the produced answer (extracts 3a, 11a), projecting some kind of trouble in the item from the very beginning, (2) their initiation of a pre-expansion sequence establishing the necessary pre-conditions and laying the groundwork for the statement of the problem (Schegloff, 2007; Stivers, 2013) (extracts 2a, 4a, 5a, 9a, 10a, 11a, 12a), (3) their use of a preface projecting the announcement of the problem (extracts 1a, 4a, 5a, 7a, 8a), and/or (4) their use of a parenthetical insert providing additional information about the answer expected for the language test item or about the problem that is to be announced (extracts 1a, 6a), which establishes shared understanding and common ground among the group before proceeding to the base sequence.

As Table 6 shows, in the pre-expansion sequence, it is found that the reviewer who problematizes the item generally initiates a preliminary inquiry (Pomerantz & Heritage, 2013) into the problem by addressing a question to the other participants about the matter at hand with reference to past instructional events based on a shared curriculum (Can Daşkın & Hatipoğlu, 2019; You, 2015) (extracts 2a, 12a) or shared sources as external sources of authority (Jacknick, 2013) (e.g., the story book, extract 10a) in search of concrete evidence for the problem, by soliciting confirmation for the answer of the item, or by expressing uncertainty about the correct answer. These interactional resources in the preliminary inquiry project a problem in the item and help perform such actions as; (1) a process of collaborative inquiry about the issue raised with respect to the item (e.g., orientation to the test material to read the other possible option mentioned or the textbook to check where the tested word "*ambitious*" appeared), (2) the identification of the sought-after evidence for the projected problem (e.g., finding out that the verb "*cover*" in the item was not studied in class), (3) the presentation of account and evidence for the expected correct answer (e.g., showing the part of the reading text where the students are expected to derive the answer from or making reference to past instructional events and textbooks where the word key to the correct answer is assumed to have been studied), (4) the tester's pre-emptive explanation of the item writing process and his/her aim in the design of the item (e.g., explaining the intended meaning in one of the options), (5) the tester's announcement of the expected answer, and (6) the consideration of the possibilities that could pose a threat to the validity of the expected correct answer (e.g., questioning whether the students would know the adjective "*efficient*" that is key to the correct answer).

These actions show that the participants give careful thought to the item in the pre-expansion sequence, which allows the problematizer of the item to confirm the problem and ensure shared understanding and common ground on which to base the problem. In other words, the preliminary inquiry encourages the production of the projected problem as the participants' responses act as a go-ahead response (Schegloff, 2007) and suggests an upcoming acceptance of the problem likely (Pomerantz & Heritage, 2013). In this respect, the preliminary inquiry serves the base sequence of problematizing the item by establishing the necessary pre-conditions and laying the groundwork (Schegloff, 2007; Stivers, 2013). In addition, it invites a collaborative investigation and identification of the problem and manages to involve all the co-participants in the process of problematizing the item in and through the talk-in-interaction. Consequently, they collaboratively show an orientation to the testing principle that was violated.

**Table 6. Interactional Resources for Initiating Preliminary Inquiry into the Problem**

| No | Preliminary Inquiry                                                             | Extract No   | Examples                                                                                                                                                                                                 |
|----|---------------------------------------------------------------------------------|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Asking a question with reference to past instructional events or shared sources | 2a, 10a, 12a | hatırlıyor musun derslerde böyle hani cover'ın (.) past tense'i geçti mi?/<br><i>do you remember if ever in classes like you know the past tense of "cover" was encountered?</i> (lines 1-3, extract 2a) |
|    |                                                                                 |              | ambitious: (.) er: nerde geçiyodu:/<br><i>where did "ambitious" appear?</i> (line 1, extract 12a)                                                                                                        |
| 2  | Requesting confirmation for the answer of the item                              | 4a, 5a       | dört "d" di mi/<br><i>four is "d" isn't it?</i> (line 2, extract 4a)                                                                                                                                     |
| 3  | Expressing uncertainty about the correct answer                                 | 9a           | altı "c" mi "d" mi bilemedim /<br><i>I'm not sure if six is "c" or "d"</i> (line 1)                                                                                                                      |

As a predominant interactional action emerging in the phase of PI, the projection of the announcement of the problem identified in the item is realized through some other interactional resources that appear in the preface or pre-expansion sequence besides the preliminary inquiry (See Table 7). As illustrated in Table 7.1, the findings demonstrate that



the reviewers project an upcoming negative evaluation of the item and an opposition to it mostly by using a verbal statement, which draws the other participants' attention to the item. In one of those verbal statements as in extract 8a, the reviewer presents the projected problem on behalf of all the teachers teaching the same group of students and raising the possibility of their objection to the item (i.e., >şimdi buna< hocalardan şöyle bi iti↑raz gelebilir/*now an objection such as this might come to this from the teachers*), which creates a more convincing context for problematizing the item by projecting the problem as the product of shared understanding among teachers. In addition, in some cases in the analysed extracts, the verbal preface is found to be embodied with a non-verbal behaviour such as pointing at the problematic word in the item or in the reading text (extracts 5a, 8a), which projects the cause of the problem and thus, prompts the other participants to anticipate the projected problem and provide a pre-emptive account for it as the tester does in extract 5a, leading to further negotiation and joint evaluation of the item from different perspectives.

The other interactional resources used to project the announcement of a problem in the item involve (1) prefacing the turn of the statement of the problem by a *but*-conjunction, which expresses contrast (extracts 1a, 7a), (2) providing background information about the emergence of the problem in the item (extracts 5a, 10a), which leads to the tester's candidate evaluation of the item as in extract 5a and her pre-emptive account for the expected correct answer as in extract 10a, and (3) requesting a change in the part of the item that is projected to have a problem (extract 1a), which gives the other participants some time to think about the item as they are all fully occupied with the test material.

**Table 7. Interactional Resources for Projecting the Problem in the Item**

| No | Projection of the Problem         | Extract No | Examples                                                                                                                                                                                                                                                                                                                                                                                                                    |
|----|-----------------------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | A verbal statement                | 4a, 5a, 8a | <p>&lt;ikide&gt; benim şöyle bi yorumum vardı/<br/> <i>for two I have a comment such as this</i> (lines 1-2, extract 5a)</p> <p>&gt;şimdi buna&lt; hocalardan şöyle bi iti↑raz gelebilir/<br/> <i>now an objection such as this might come to this from the teachers</i> (line 1, extract 8a)</p>                                                                                                                           |
| 2  | But-prefaced turn                 | 1a, 7a     | <p>ama might riskli mi diye bi düşünüyorum/<br/> <i>but I am thinking if "might" is risky</i> (line 4, extract 7a)</p>                                                                                                                                                                                                                                                                                                      |
| 3  | Presenting background information | 5a, 10a    | <p>ya <u>biz</u> belki bildiğimiz için bize öyle geldi bilemiyorum/<br/> <i>you see, maybe it is because we know it that it seemed that way to us, I don't know</i> (lines 12-13, extract 5a)</p> <p>hani <u>ben</u> (.) &gt;açıkcası hiç&lt; kardeşlerinin karak(.)te-terlerine çok takılmadım /<br/> <i>you know frankly speaking I never focused much on the characters of his siblings</i> (lines 7-9, extract 10a)</p> |
| 4  | Request for a change in the item  | 1a         | <p>&gt;bunun yerine başka bişey düşünebilir&lt;<u>miyiz</u> dedim/<br/> <i>I thought whether we could think of something else for this</i> (lines 8-9)</p>                                                                                                                                                                                                                                                                  |

Following the preliminary work which not only projects a problem in the item but also establishes the necessary pre-conditions for its announcement, the base FPP of PI is then launched where the problem is overtly announced and explained with evidence as the third step in Figure 3. It is found that the interactional act of giving account and evidence for the problem is an inevitable part of PI, appearing in all the extracts analysed for this study. As summarized in Table 8, the reviewers make an attempt to prove the problem by presenting evidence from the available shared sources such as the reading text, textbook, listening notes/script, and the story book as external sources of authority (Jacknick, 2013) (e.g., çünkü beyin (.) fiziksel reaksiyonlar yapıyo/*because the brain*

*makes physical reactions*, extract 6a) as well as the past instructional events (Can Daşkın & Hatipoğlu, 2019; You, 2015) (e.g., *öyle bişey geçmedi* ↑*de/such a thing did not appear*, extract 2a) and the students' present state of L2 knowledge and skills (e.g., *under* prefix'iyle ilgili bi bilgi sahibi değil/*they have no knowledge about the prefix "under"*, extract 8a). While providing account, they are also found to translate the key words or sentences in the item and in the relevant parts of the texts into their L1 (i.e., Turkish) (extracts 4a, 7a, 9a) in order to make a connection between the item and the presented evidence from, for instance, the reading text and to clarify their point. In some cases as in extracts 2a, 10a and 12a, the reviewers announce the problem based on the evidence they have obtained through their preliminary inquiry in the pre-expansion sequence (e.g., *çok onu bulamayabilirler*/*they might not be able to find it*, extract 2a).

While providing a justification for the problem in the item, the reviewers at the same time show an orientation to the testing principles for item writing. To illustrate, in extract 1a, BA refers to the implausibility of the correct option with her account "*çok (.) böyle (0.4) açık olmuş*/*it has become too like obvious*" and in extract 3a, the reviewers DE and BA refer to the violation of "only one correct answer" principle as they present evidence from the text (e.g., *ikisi de oluyo?*/*both are possible*). As shown in Table 8, the analysis has revealed that one of the major violated testing principles that the participants make reference to while giving account for the problem in phase PI is the existence of more than one possible correct answer in an item (extracts 3a, 4a, 6a, 7a, 9a, 11a) followed by testing what you teach principle (extracts 2a, 8a, 12a). They also orient to the violation of testing principles related to the use of (1) unnatural language in the item (e.g., the use of "*before*" as a prompt for past perfect tense in extract 5a), (2) complex vocabulary in a reading text (i.e., "*underrecognize*" in extract 8a), (3) ambiguous terms in the item ("*can*" in extract 9a) as well as the content that is tested in the item (extract 10a), the independency of the individual items and the difficulty level of the item as in extract 12a.

**Table 8. Summary of Accounts and Evidence Presented for the Problems in the Items**

| Extract No | Item Type                                                                                                                                    | Account/Evidence for the Problem                                                                                                                                                                                                                                                                                                                                                              |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1a         | A multiple-choice reading comprehension item that asks what is <i>not</i> one of the advantages of the tribe Kalenjini mentioned in the text | The text being about a tribe makes the mention of “ <i>experts</i> ” in the correct option too obvious to identify for the students.<br>(Plausibility principle)                                                                                                                                                                                                                              |
| 2a         | A fill-in-the blanks item (tense passage)                                                                                                    | The verb “ <i>cover</i> ” to be transformed into its past tense form is not taught in class.<br>(Test what you teach principle)                                                                                                                                                                                                                                                               |
| 3a         | A multiple-choice cloze test item where the expected correct answer is “ <i>furthermore</i> ”                                                | The target sentence can be regarded as a consequence of the action explained in the previous sentence (i.e., <i>they have broken off from much larger areas of ice...</i> ) and thus, the distractor “ <i>therefore</i> ” can also be correct.<br>(Only one correct answer principle)                                                                                                         |
| 4a         | A multiple-choice reading comprehension item                                                                                                 | The distractor option “ <i>customers do shopping very fast</i> ” can also be correct since the stem is designed in general terms (i.e., <i>when muzak soundscapes are used</i> ) and the information in the option is mentioned in the fifth paragraph in the text.<br>(Only one correct answer principle)                                                                                    |
| 5a         | A multiple-choice cloze test item where the expected correct answer is the past perfect tense “ <i>had seen</i> ”                            | The time expression “ <i>before</i> ” in the target sentence given as a prompt for past perfect tense is inappropriate and unnecessary within the given context.<br>(Appropriacy of language principle)                                                                                                                                                                                       |
| 6a         | A fill-in-the-blanks item (lecture outline completion) where the expected answer is “ <i>chemicals</i> ”                                     | The alternative “ <i>physical reactions</i> ” mentioned in the lecture can also be correct considering the notes presented in the item about the physical effects of fear (i.e., <i>the brain, making -blank-</i> ) since it is explained in the lecture that the brain makes chemicals that cause physical reactions such as heart beat and sweating.<br>(Only one correct answer principle) |
| 7a         | A binary-choice cloze test item where the expected answer is “ <i>should</i> ”                                                               | The distractor modal verb “ <i>might</i> ” can also be correct within the context given in the item where the verb expressions “ <i>be careful</i> ” and “ <i>protect</i> ” can be interpreted as a possibility.<br>(Only one correct answer principle)                                                                                                                                       |

**Table 8 (continued)**

|     |                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 8a  | A multiple-choice reading comprehension item                                                                                                           | The students do not know the prefix “ <i>under</i> ” and thus, will not understand the verb “ <i>underrecognize</i> ” given in the part of the text that the item is based on.<br><br>(Simplicity of vocabulary principle, Test what you teach principle)                                                                                                                                                                                                                                               |
| 9a  | A multiple-choice reading comprehension item, asking about the opinion of the famous striker Michael Owen’s father with regard to his son’s profession | The distractor option “ <i>people can freely criticize Michael when he does not play</i> ” can also be correct when the students interpret the meaning of “ <i>can</i> ” in the option as a possibility rather than a permission based on the sentence “ <i>according to him the worst parts of being famous are the bad public reactions which appear when the coach sends Michael off the field</i> ” given in the text.<br><br>(Only one correct answer principle, Avoidance of ambiguity principle) |
| 10a | A matching test item in a story exam where the students are expected to match the character descriptions with the names given in a box                 | The expected answer “Lizzie” is one of the minor characters in the story that the reviewer did not focus much on while reading the story.<br><br>(Avoidance of trivial content principle)                                                                                                                                                                                                                                                                                                               |
| 11a | A multiple-choice reading comprehension item                                                                                                           | Not the expected correct option “ <i>they want to appreciate the world better</i> ” but the distractor option “ <i>they want to achieve a higher status in their community</i> ” is the correct answer due to the expression “ <i>spiritual insight</i> ” given in the reading text and the unavailability of information for the expected correct answer in the reading text.<br><br>(Only one correct answer principle)                                                                               |
| 12a | A binary-choice cloze test item (vocabulary passage) where the expected answer is “ <i>ambitious</i> ”                                                 | The correct answer “ <i>ambitious</i> ” is not taught in class since personality adjectives will be explicitly studied later in the program, the target sentence where “ <i>ambitious</i> ” is used is difficult to understand, and two points are tested subsequently in one single sentence.<br><br>(Test what you teach principle, Level of difficulty principle, Independency principle)                                                                                                            |

The findings illustrate that the reviewers generally mitigate and downgrade the disaffiliative impact of the problem they have identified in the language test item by using some interactional resources (See Table 9) while announcing it in the base FPP, or projecting it in the pre-expansion sequence or the preface. First of all, before making a justification for the problem, some of the reviewers are seen to present the problem through a direct (*onu da kabul eder miyiz?/would we accept that too?*, extract 6a), or an embedded question (*might riskli mi diye bi düşünüyorum/I am thinking if "might" is risky*, extract 7a) for everyone in the IR group to think about. This paves the way for a collaborative search of evidence for the problem and collaborative inquiry into the raised issue as the participants orient to the test material. In this way, the problematizers mitigate the disaffiliative impact of the announced problem in the item by leaving it to the other participants to decide and preparing them for the upcoming explanation that the questions entail.

Secondly, the reviewers are seen to use some hedges (e.g., *sanki/as if*, *belki/maybe*, *galiba/I think*, *gibi/like*), which express possibility and treat the identified problem as one possible issue for the group to consider before approving the item. Thirdly, as in extracts 5a and 10a, the reviewers provide a mitigating account for their upcoming negative evaluations of the item by giving background information about why the item has caused a problem for them (See Table 9.3), which also projects the problem in the item as explained in Table 7.3. For instance, in extract 5a, the reviewer MI prefaces her announcement of the problem by explaining that the reviewers' realization and shared knowledge that "*before*" was added to the target sentence by the tester might have caused them to see it as a problem. Similarly, in extract 10a, in the pre-expansion sequence, YU provides an account for his failure to match one of the descriptions in the item with a character name based on his own personal way of reading the story book as evident in his use of an emphasized first person pronoun. In this way, the reviewers downgrade their upcoming claim for the item by showing that it is their own way of approaching the item that might have caused the problem and leaving it to the tester to decide if it could be considered as a serious problem.

Finally, as in extract 10a, it is seen that offering a positive assessment is another interactional resource used to mitigate the problem that is to be announced with respect to

the item. The reviewer YU, for example, does not directly address the problematized item but first offers a positive assessment of the other three character names he could match in the story exam in an attempt to mitigate the disaffiliative impact of his upcoming negative evaluation with regard to the expected answer “Lizzie” in one of the items.

**Table 9. Interactional Resources for Mitigating the Problem in the Item**

| No | Mitigating the Problem      | Extract No | Examples                                                                                                                                                                                                                                                                        |
|----|-----------------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Direct or embedded question | 4a, 6a, 7a | <p>make physical reactions onu da kabul eder miyiz?/<br/> <i>would we accept that too?</i> (lines 7-9, extract 6a)</p> <p>might riskli mi diye bi düşünüyorum /<br/> <i>I am thinking if "might" is risky</i> (line 4, extract 7a)</p>                                          |
| 2  | Hedges                      | 5a, 12a    | <p>belki / <i>maybe</i> (line 12, extract 5a)<br/> bilemiyorum / <i>I don't know</i> (line 13, extract 5a)<br/> sanki / <i>I think, as if</i> (lines 9, 27, extract 12a)<br/> galiba / <i>I think</i> (line 19, extract 12a)<br/> gibi / <i>like</i> (line 27, extract 12a)</p> |
| 3  | Mitigating account          | 5a, 10a    | <p>ya <u>biz</u> belki bildiğimiz için bize öyle geldi bilemiyorum /<br/> <i>you see, maybe it is because we know it that it seemed that way to us, I don't know</i> (lines 12-13, extract 5a)</p>                                                                              |
| 4  | Positive assessment         | 10a        | <p>tom (.) carver ve gwenny' <math>\uparrow</math>yi beğendim burda: /<br/> <i>I liked tom, carver and gwenny here</i> (line 3)</p>                                                                                                                                             |

The announcement of the problem makes agreement or disagreement relevant next in the base SPP as the sequentially-fitted next response, which is the fourth step in Figure 3. As a result, the group either directly agrees with the problem in the base SPP (extracts 1a, 2a, 5a, 6a, 7a, 8a, 12a) (labelled as “yes” in Figure 3) or some kind of resistance is displayed with regard to the announced problem (extracts 3a, 4a, 9a, 10a, 11a, 12a) before reaching a consensus on the raised issue (labelled as “no” in Figure 3). Though not always the case, the

base SPP of displaying a stance of agreement or disagreement may then be followed by the participants' initiation of a non-minimal post-expansion sequence (NMP) inviting further elaboration of the problem (Schegloff, 2007; Stivers, 2013) as in extracts 3a, 4a, 6a, and 9a.

When the group agrees with the problem, the sequential pattern is seen to unfold in three alternative ways before the phase ends. Firstly, in most of the cases, the phase of SC is projected as the relevant next action to overcome the problem (extracts 1a, 2a, 5a, 7a, 8a, 9a, 10a, 12a), which signals continuity in the IR interaction. It is commonly seen that the participants request and seek a suggestion for a change in the item (e.g., *oraya başka bişey ne olabilir/what else can be put there*, extract 7a), which not only projects continuity in the IR interaction but also signals the participants' willingness to change the item and move onto the next phase of SC in the IR interaction. In this way, the participants also indirectly convey their agreement with the problem in the base SPP as in extracts 5a, 8a, 10a, 12a. In extracts 5a and 7a, the participants' projection of the phase of SC as the relevant next action is also embodied as they orient to the test material and other relevant sources (e.g., textbook) in search of a suggestion for a change in the item to overcome the problem and improve the item. Secondly, following the participants' agreement with the problem, the IR group can decide to omit the item as in extract 12a since the identified problem(s) leave no room for change and improvement in the item. Thirdly, although the participants' agreement with the problem does not project an expansion at the PI stage, the IR group may initiate a non-minimal post-expansion sequence for further elaboration of the problem as in extracts 4a and 6a.

On the other hand, a participant might challenge and display disagreement with the problem or a dispute might arise in the group over the articulated problem in the base SPP (i.e., “no” in Figure 3). In doing this, the participants challenge a claim by presenting counter-evidence from the available shared sources (e.g., the reading text, story book) and making counter-claims (See Table 10). For instance, in extract 9a, the participants initiate a process of negotiation of meaning and continue to challenge each others' claims and accounts in the non-minimal post-expansion sequence (e.g., *aslinda (.) worst parts diyo* *↑ya/actually you know it says “worst parts”*) until an agreement is reached on the problem. Orienting to the reading text, the story book or the item itself, the testers often provide counter-evidence that challenges the problem (e.g., *aslinda burda*



şey var/*actually here there is this thing* uhm in a store (.) people seem to shop in a more relaxed and leisurely manner, extract 4a). In extract 9a, the tester SA also explains her aim in the design of the problematized distractor by clarifying what she actually meant by her use of “can” (e.g., *aslında burda can'den benim (.) er amacım şeydi/actually here with “can” my aim was the thing*). In this way, she challenges the problem by referring to the item writing process. Additionally, while challenging a problem, the testers make a claim that rejects the violation of the testing principles addressed by the problematizers of the item. To illustrate, in extract 3a, ME disagrees that the item has three correct answers by claiming that one of the problematized distractors “however” cannot be possible as a correct answer (i.e., *however olmaz hocam/“however” is not possible miss*). Similarly, in extract 12a, SA rejects the violation of the testing principle that “the content of individual items should be independent of each other” (Haladyna & Rodriguez, 2013) by explaining that the sentence that is claimed to test two points subsequently is connected with “but” and implying that the two items do not overlap (i.e., *en azından but'la bağlanmış/at least it is connected with “but”*). Through those turns of challenging the problem, the testers state the underlying reason for their construction of the item in such a way and try to prove that the items are in line with the testing principles.

Considering the disagreement formulations in the analysed extracts, it is seen that in extracts 4a, 9a, and 10a, the testers launch a “pro forma” agreement (Schegloff, 2007) (i.e., an agreement-prefaced disagreement) whereby a minimal agreement that is initially displayed either in the base SPP or in the non-minimal post-expansion sequence (extract 9a) then turns into a disagreement (i.e., arrow moving from “yes” to “no” in Figure 3) as the participants exhibit partial resistance by providing counter-evidence for the problem. Their “pro forma” agreement (Schegloff, 2007) serves to delay their upcoming dispreferred response (i.e., challenging of the problem illustrated with “no” in Figure 3) as a contiguity-breaking practice and thus, to mitigate the disaffiliative impact of their projected display of disagreement (e.g., *işte anladım hocam ben orda can'i şey diye:- (.) aslında evet can'i koymamın hedefi/you see I understand miss, there I put “can” so as to- actually yes, my aim for putting “can” is-*, extract 4a).

**Table 10. Summary of Challenges Presented to the Announced Problems in the Items and Their Outcomes**

| Extract No | Challenge                                                                                                                                                                                      | Outcome                                                                                                                                                                |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3a         | One of the two distractors (i.e., <i>however</i> ) initially claimed to be correct is actually not possible as a correct answer within the given context.                                      | Challenge accepted / Problem redesigned and only one distractor (i.e., <i>therefore</i> ) is regarded as a possible correct answer.                                    |
| 4a         | There is information in one part of the reading text that makes the distractor option incorrect.                                                                                               | Challenge rejected / Problem accepted due to the availability of some other information in the text making the option correct.                                         |
| 9a         | There is an expression in the reading text (i.e., <i>the worst parts of being famous</i> ) that makes the distractor option incorrect.                                                         | Challenge rejected / Problem accepted due to the students' possible interpretation of " <i>can</i> " in the distractor option as a possibility rather than permission. |
| 10a        | The character Lizzie that is claimed to be a minor detail is mentioned twice in the story and even at the very end of the story, some information is given about her.                          | Challenge rejected / Problem accepted                                                                                                                                  |
| 11a        | The expression " <i>in their community</i> " in the distractor option makes it incorrect and the seventh paragraph in the text provides enough evidence for the expected correct answer.       | Challenge accepted / Problem rejected / Item unchanged                                                                                                                 |
| 12a        | The target sentence in the binary-choice cloze test that is claimed to test two points is at least connected with " <i>but</i> " and can be broken into two sentences with " <i>however</i> ". | Challenge accepted / Item omitted due to the emergence of other problems in the item                                                                                   |

The findings demonstrate that such an interactional act of challenging the announced problem eventually results in a display of mutual understanding and affiliation with the participants' stance and the resolution of the emerging dispute in and through the talk-in-interaction. This is achieved through (1) further account-giving and elaboration of the problem as a counter-challenge in response to the disagreeing participant in a non-minimal post-expansion sequence (extracts 3a, 9a) as illustrated in Figure 3, (2) the problematizer's redesign of a prior account from a different perspective (e.g., the students' ability) to elicit a preferred response (i.e., agreement with the problem) (i.e., *çocuk ama <onu o:rdan> (.) çıkartamaz gibi/but it is like the kid cannot derive it*

from there, extract 9a), (3) the problematizer's alignment with the disagreeing participants by downgrading her competence in answering the question (i.e., claiming carelessness and lack of attention) upon the group's discussion of the test item through accounts and (counter)-challenges (extract 11a), (4) the disagreeing participant's quick recognition and understanding of the problem before the other participants' counter-challenge and her use of a 'peer referencing' (i.e., *as you said*) to turn her initial disagreement with the announced problem into agreement (ama (.) dediğiniz gibi başka paragraflarda da ona gidecek şey var/*but as you said in the other paragraphs too there is something referring to it*, extract 4a) (i.e., arrow moving from "no" to "yes" in Figure 3), (5) the disagreeing participant's backdown and modification of her position by projecting the next phase of SC in an attempt to receive a preferred response (i.e., mutual agreement in the group) and to ensure contiguity in the IR interaction as a result of no uptake and orientation to his/her counter-argument from the other participants (↑belki bire başka bişey/*maybe for one something else*, extract 10a) (Pomerantz, 1984a; Pomerantz & Heritage, 2013), where once again the disagreement that is initially displayed turns into agreement (i.e., arrow moving from "no" to "yes" in Figure 3), and (6) the identification of other kinds of problems in the item, which the participants agree with (i.e., "ambitious" not being taught in class and the target sentence being difficult to understand, extract 12a).

As a result, a consensus is reached in the IR groups although different decisions are seen to be made with regard to the item. As Table 10 shows, the groups generally agree with the announced problem(s) and decide to revise the item as in extracts 3a, 4a, 9a, 10a or to omit the item from the test as in extract 12a. Additionally, in extract 3a, further account-giving and rechecking of the problem entailed by the tester's act of challenging the problem leads to the redesign and redefinition of the already announced problem. Therefore, the group decides that not two but only one of the other distractor options in the item (i.e., *therefore*) can be possible as a correct answer. In another case as in extract 11a, the act of challenging this time disproves the problem claimed for the item and thus, the group decides to keep the item unchanged after agreeing that it, in fact, causes no problem through a process of considerable discussion on the item. In this instance, as a result of an emerging dispute with regard to the correct answer of the item, different from the other extracts, not the problematizer of the item but the participants who challenge the problematizer through

account and evidence giving turn out to be right in their argument for the validity of the item and hence, the group eventually approves the item.

The findings demonstrate that throughout the phase of PI, the participants often go beyond the use of an agreement token and display mutual agreement through a collaborative process. As Table 11 shows, they achieve this by (1) using a full or a partial modified repeat (Kim, 2002; Stivers, 2005) of a claim made about the item, (2) offering a (re)formulation of the announced problem (Heritage & Watson, 1979), (3) presenting additional account or evidence that elaborates on the problem with an orientation to the test material and other relevant sources (e.g., tapescript) as external sources of authority (Jacknick, 2013), (4) offering a candidate understanding (Antaki, 2012; Heritage, 1984b) of the presented problem or an account, (5) employing an anticipatory completion (Lerner, 1996a), which may appear in the form of a candidate understanding of the problem (extracts 2a, 6a) or in the form of a candidate evaluation of the item (extract 1a), (6) engaging in a choral co-production (Lerner, 2002) with the other participants while giving account, (7) collaboratively completing each others' turns while presenting evidence from the reading/listening text, (8) the tester reflecting on the item writing process for further elaboration of the problem, (9) offering a candidate evaluation of the item, and (10) providing a candidate explanation for the problem to make the already presented account more specific.

These interactional resources enable the participants to demonstrate their understanding of the matter at hand and affiliate with the stance taken by the problematizer of the item (Hayashi, 2013). They also overtly confirm the announced problem and further convey alignment and affiliation. Based on this, by means of such interactional resources, the participants further elaborate on the problem especially in a non-minimal post-expansion sequence, which is predominantly characterized by collaborative account and evidence giving from various perspectives (extracts 3a, 4a, 6a, 9a). In this way, the participants clarify their prior claims with regard to the item and reassert the problem even when an agreement is displayed in the base SPP, which does not project an expansion in phase PI as in extracts 4a and 6a. As a result, they establish mutual understanding about the item in the IR group and make sure that they have given the item a careful and a critical thought before proceeding to the next phase of the interaction.

**Table 11. Interactional Resources for Displaying Mutual Agreement through a Collaborative Process in Phase PI**

| No | Display of Mutual Agreement                 | Extract No                   | Examples                                                                                                                                                                                                                               |
|----|---------------------------------------------|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Full or partial modified repeat of a claim  | 1a, 3a, 4a, 5a, 6a, 11a, 12a | böyle bişey yok/ <i>there is no such thing</i> (line 7, extract 1a)                                                                                                                                                                    |
| 2  | (Re)formulation                             | 3a, 4a, 6a, 7a, 9a, 10a, 11a | oraya da gidebilir çocuk evet/ <i>yes the kid can also go for it</i> (line 51, extract 4a)<br>o kısım evet karşılıyo onu/ <i>that part yes corresponds to that</i> (line 31, extract 9a)                                               |
| 3  | Additional account/evidence                 | 3a, 6a, 9a, 11a              | zaten aslında öyle bir (0.2) er: ifade var orda şöyle diyo<br>your brain quickly makes chemicals that cause different physical reactions/ <i>actually there is already such an expression there, it says</i> (lines 58-62, extract 6a) |
| 4  | Candidate understanding                     | 1a, 2a, 8a, 11a              | kelimeyi mi değiştirelim/ <i>is it the word that we shall change</i> (line 5, extract 8a)<br>bütün bir paragraf olarak mı/ <i>is it the paragraph as a whole</i> (line 87, extract 11a)                                                |
| 5  | Anticipatory completion                     | 1a, 2a, 5a, 6a               | olmuyacağı çok açık/ <i>it is too obvious that it is not the one</i> (line 15, extract 1a)                                                                                                                                             |
| 6  | Choral co-production                        | 7a, 9a, 11a                  | olabilirler/ <i>might be</i> (lines 10-11, extract 7a)                                                                                                                                                                                 |
| 7  | Collaborative turn completions              | 6a, 9a                       | lines 43-48 (extract 6a)                                                                                                                                                                                                               |
| 8  | Tester's reflection on item writing process | 5a, 6a                       | orda aslında bişey vardı da ben uzatmadım/ <i>there was actually something there but I didn't extend it</i> (lines 30, 33, extract 6a)                                                                                                 |
| 9  | Candidate evaluation of the item            | 1a, 5a                       | çok mu net dicesiniz/ <i>you will say that it is too obvious</i> (line 14, extract 5a)                                                                                                                                                 |
| 10 | Candidate explanation for the problem       | 3a                           | broken off'dan dolayı (.) there(.)fore'a gittiniz di mi/ <i>because of "broken off" you chose "therefore" didn't you?</i> (lines 26, 28)                                                                                               |

The findings show that the talk-in-interaction in phase PI displays an orientation to the *institutional role distribution of the participants*. As revealed in this study, the interactional actions emerging in the phase of PI show that the testers and reviewers have certain roles and responsibilities in this phase. To begin with the interactional indicators of tester roles and responsibilities, as illustrated in Table 12, the tester's interactional acts of (1) announcing the expected correct answer or (dis)confirming the reviewers' answers to the test items, (2) providing an elaborative account for the expected correct answer to provide insight into the projected or announced problem, (3) challenging the problem where necessary by giving counter-evidence and making counter-claims, and (4) explaining the item writing process and her aim in the item as well as the problematizer soliciting the tester's confirmation for the correct answer (extracts 4a, 5a, 11a) and his/her correct recall of information from the story book (extract 10a) all signal the tester's epistemic right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item (Type 1 knowledge, Pomerantz, 1980) as the writer and thus, the 'owner' of the item. Therefore, those interactional actions position the testers as already knowing (K+) recipients (Heritage, 2012, 2013) and ascribe them an epistemic authority.

On the other hand, in some cases where the announced problem concerns the teaching domain as in extracts 2a, 8a, 12a, the tester is ascribed a K- epistemic status (Heritage, 2013) as the emerging problem in the item presents an issue that she might not have foreseen while constructing the item. For instance, in extract 2a, the matter at hand concerns the teaching practice and the students' knowledge of a word (i.e., *cover*) given in the item. Since it is a domain of information that the tester does not have an equal epistemic access to with the teachers, the problematizer of the item asks the other reviewer, not the tester about the issue (i.e., the teaching of the verb "*cover*") and only after verifying the problem, she turns to the tester to announce it, ascribing a K- epistemic status to her. For the same reason, as the reviewers problematize the item, they shift their gazes to the tester in particular (extracts 1a, 2a, 3a, 4a, 6a, 8a, 9a, 11a, 12a) and explicitly address her with her name or an endearment address term (extracts 2a, 4a, 12a). In this way, they orient to the tester's right to learn about the problem as the writer of the item, projecting an unknowing status for her since they raise an issue about the item that might have gone unnoticed by the tester or that is about the teaching context which the tester has a relative lack of epistemic access to.

**Table 12. Interactional Indicators for Tester Roles and Responsibilities in Phase PI**

| No | Tester Roles and Responsibilities                                                                  | Extract No                       | Examples                                                                                                                                                                                                                                                    |
|----|----------------------------------------------------------------------------------------------------|----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Announcing the expected correct answer or (dis)confirming the reviewers' answers to the test items | 3a, 4a, 5a, 6a, 7a, 9a, 10a, 11a | onun cevabı lizzie/ <i>the answer to that is lizzie</i> (line 15, extract 10a)<br>denizliyi hocam<br>hedeflemiştim/ <i>I aimed for "d" miss</i> (line 5, extract 4a)                                                                                        |
| 2  | Providing an elaborative account for the expected correct answer                                   | 1a, 10a, 11a                     | lizzie'den iki kere bahsediyο/ <i>it mentions lizzie twice</i> (line 14, extract 10a)                                                                                                                                                                       |
| 3  | Challenging the announced problem                                                                  | 3a, 4a, 9a, 10a, 11a, 12a        | however olmaz hocam/ <i>"however" is not possible miss</i> (line 22, extract 3a)<br>aslında burda şey var uhm in a store (.) people seem to shop in a more relaxed and leisurely manner/ <i>actually here there is this thing</i> (lines 38-41, extract 4a) |
| 4  | Explaining the item writing process and her aim in the item                                        | 5a, 6a, 9a                       | orda aslında bişey vardı da ben <u>uzatmadım</u> / <i>there was actually something there but I didn't extend it</i> (lines 30, 33, extract 6a)                                                                                                              |
| 5  | Requesting confirmation for her understanding of the problem                                       | 1a, 3a, 8a                       | kelimeyi mi değiştirelim/ <i>is it the word that we shall change</i> (line 5, extract 8a)                                                                                                                                                                   |
| 6  | Being alert to upcoming problems in the item                                                       | 1a, 5a, 10a                      | çok mu net dicesiniz/ <i>you will say that it is too obvious</i> (line 14, extract 5a)<br>şunu zaten- (.) er birinciyi yapamamışsın/ <i>that after all you could not do the first one</i> (line 1, extract 10a)                                             |

Furthermore, the testers' confirmation requests (See Table 12.5) in the form of a candidate explanation for (extract 3a) or a candidate understanding (Antaki, 2012; Heritage, 1984b) of the problem (extracts 1a, 8a) orient to their role to check their understanding of the problem as the writer of the item who is accountable for the test. Their correct understanding of the problem identified in the test item is important for accountability issues in testing because in any case of criticism, as the tester of the institution, they will have to explain to the stakeholders why the item is designed in such a way. In addition, as illustrated in Table 12.6,

an orientation is displayed to the testers' role and responsibility to be alert to any problems that might be addressed in the IR interactions with respect to language test items. This is evident in the testers' candidate evaluations of the item (extracts 1a, 5a) as a result of their anticipation of the projected problem through, for example, the problematizer's non-verbal behaviour of pointing at the problematized word in the item in extract 5a. Similarly, in extract 10a, the tester initiates a pre-expansion sequence upon her quick observation that the item was left unanswered by the reviewer, which demonstrates her readiness to understand and act on any problem that might arise in the item.

As for the reviewer roles, as Table 13 demonstrates, throughout the phase of PI in the IR interactions, the reviewers show an orientation to the responsibilities that they should carry out both before and during the IR session. In the former case (i.e., *before* the IR session), the reviewer responsibilities that are oriented to in the talk-in-interaction involve (1) taking the test as a student would before giving feedback to the tester in the IR interaction, which is evident in the participants' reference to their test-taking process by using the past tense or present perfect tense (e.g., *işaretlemişiz/we marked*, extract 3a; *yapamamışsın/you could not do*, *takılmadım/did not focus*, extract 10a; *yanlış mı yapmışım/have I done it wrong?*, *ben "a" demişim/I have said "a"*, *algıladım/perceived*, *oraya bakmamışım/I have not looked there*, extract 11a), and (2) reading the story carefully before coming to the IR session for the story exam as illustrated by the reviewer YU's reference to his reading process in extract 10a (i.e., *kardeşlerinin karak(.)te-terlerine çok takılmadım/I did not focus much on the characters of his siblings*).

Regarding the reviewer roles and responsibilities to be carried out *during* the IR interactions, the reviewers display an orientation not only to their roles as problematizers but also as testers and teachers. Firstly, as exemplified in Table 13, their role as problematizers of the language test items that they have reviewed is made evident in their (1) preliminary inquiry about the problem before announcing it, (2) projection of the announcement of the problem that they have identified in the item, (3) account and evidence while stating the problem in the item, (4) mitigation of the disaffiliative impact of their negative evaluation of the item, and their (5) collaborative response to the challenges presented by the tester with respect to



the problem identified in the item. In addition, in cases where there are two reviewers, although one of them usually takes an initiative to announce the problem in the item, the reviewers convey a similar stance towards the item and do not challenge each other, further orienting to their role as the problematizers of the item. This is the reason why the reviewers use an inclusive language for their approach to the item in extracts 3a (i.e., *ikimiz de aynı şeyi işaretlemiştir/we both marked the same thing*) and 5a (i.e., *biz belki bildiğimiz için bize öyle geldi/maybe it is because we know it that it seemed that way to us*), which shows that they share the responsibility for problematizing the item and put themselves in equal epistemic position. As a result, they, for instance, happen to counter-challenge the tester collaboratively by giving account and evidence for the problem. However, extract 11a is an exception where the tester and one of the reviewers this time adopt the same stance towards the item, challenging the other reviewer DE since the item actually has no problem but a dispute arises due to DE's incorrect answer as a result of her carelessness and lack of attention to the part of the reading text that has led to the expected answer.

**Table 13. Interactional Indicators for Reviewer Roles and Responsibilities in Phase PI**

| No                                          | Reviewer Roles and Responsibilities                   | Extract No                    | Examples                                                                                                                                                                                              |
|---------------------------------------------|-------------------------------------------------------|-------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Before the IR Session</b>                |                                                       |                               |                                                                                                                                                                                                       |
| 1                                           | Taking the test as a student                          | 3a, 10a, 11a                  | <i>ikimiz de aynı şeyi işaretlemiştir/we both marked the same thing</i> (lines 2-4, extract 3a)                                                                                                       |
| 2                                           | Reading the story for the reviewing of the story exam | 10a                           | <i>kardeşlerinin karak(.)te-terlerine çok takılmadım/ I did not focus much on the characters of his siblings</i> (lines 7-9)                                                                          |
| <b>During the IR Session</b>                |                                                       |                               |                                                                                                                                                                                                       |
| <i>Reviewers as Problematizers of Items</i> |                                                       |                               |                                                                                                                                                                                                       |
| 3                                           | Carrying out preliminary inquiry about the problem    | 2a, 4a, 5a, 9a, 10a, 11a, 12a | <i>hatırlıyor musun derslerde böyle hani cover'ın (.) past tense'i geçti mi?/ do you remember if ever in classes like you know the past tense of "cover" was encountered?</i> (lines 1-3, extract 2a) |
| 4                                           | Projecting the problem                                | 1a, 4a, 5a, 7a, 8a, 10a       | <i>&lt;ikide&gt; benim şöyle bi yorumum vardı/ for two I have a comment such as this</i> (lines 1-2, extract 5a)                                                                                      |

**Table 13 (continued)**

|                              |                                                                              |                              |                                                                                                                                                                                                                                                                                                                      |
|------------------------------|------------------------------------------------------------------------------|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 5                            | Giving account and evidence for the problem                                  | All extracts                 | çünkü ↑zaten (0.2) biz bir er tribe gibi bişiyden bahsederken birden bire zaten expert'ten söz etmek çok (.) er: çok (.) böyle (0.4) açık olmuş/ <i>because after all as we talk about something like a tribe, talking about "expert" all of a sudden has become too, too like obvious</i> (lines 11-18, extract 1a) |
| 6                            | Mitigating the problem                                                       | 4a, 5a, 6a, 7a, 10a, 12a     | tom (.) carver ve gwenny' ↑yi beğendim burda:/ <i>I liked tom, carver and gwenny here</i> (line 3, extract 10a)                                                                                                                                                                                                      |
| 7                            | Responding to the tester's challenges about the problem                      | 3a, 4a, 9a                   | çocuk ama <onu o:rdan> (.) çıkartamaz gibi yani "d"ye de gider/ <i>but it is like the kid cannot derive it from there so s/he can also go for "d"</i> (lines 72-75, extract 9a)                                                                                                                                      |
| <i>Reviewers as Testers</i>  |                                                                              |                              |                                                                                                                                                                                                                                                                                                                      |
|                              |                                                                              |                              | böyle <u>birden</u> bire burda before fne geziyo diye düşündüm/ <i>like all of a sudden I thought what "before" is doing here</i> (lines 21-22, extract 5a)                                                                                                                                                          |
| 8                            | Orienting to testing principles                                              | All extracts                 | doğru cevap kabul edilmesi gerekir orda/ <i>it should be accepted as a correct answer there</i> (line 28, extract 6a)                                                                                                                                                                                                |
|                              |                                                                              |                              | aynı cümleden iki soru çıkarmış çı gibiyiz/ <i>it is like we have derived two questions from the same sentence</i> (lines 27-28, extract 12a)                                                                                                                                                                        |
| <i>Reviewers as Teachers</i> |                                                                              |                              |                                                                                                                                                                                                                                                                                                                      |
| 9                            | Responding to questions about instructional events or students' L2 knowledge | 2a, 12a                      | yok ((shakes her head)) öyle bişey geçmedi ↑de/ <i>no, such a thing did not appear</i> (line 5, extract 2a)                                                                                                                                                                                                          |
| 10                           | Framing claims from students', teachers' and teaching perspectives           | 2a, 4a, 6a, 8a, 9a, 11a, 12a | oraya da gidebilir çocuk evet/ <i>yes the kid can also go for it</i> (line 51, extract 4a)<br>ambitious: (.) er: nerde geçiyodu:/ <i>where did "ambitious" appear?</i> (line 1, extract 12a)                                                                                                                         |

Secondly, as seen in all the extracts, the reviewers show an orientation their institutional roles as testers by making reference to the testing principles for item writing while announcing and explaining the problem in the test item in the phase of PI. They do not just take the test as a student would and share their answers with the tester but they also critically think about each item and evaluate it in terms of the testing principles described in the literature for item writing. For instance, they review the items in terms of (1) the plausibility of the options in a multiple-choice item (e.g., çok (.) böyle (0.4) açık olmuş/*it has become too like obvious*, extract 1a), (2) their independency (e.g., aynı cümleden iki soru çıkarmış çı gibiyiz/*it is like we have derived two questions from the same sentence*, extract 12a), (3) the representation of the tested teaching point in the syllabus or the course content (e.g., sonra öğretmez galiba ambitious'ı/*I think we will teach "ambitious" later*, extract 12a), (4) the importance of the content tested (e.g., >işte< çok (.) minor bişey ya ↑o/ *you see that is something too minor*, extract 10a), (5) the language used (e.g., böyle birden bire burda before fne geziyo diye düşündüm/*like all of a sudden I thought what "before" is doing here*, extract 5a), and (6) the linguistic complexity for the proficiency level of the target group of students (bana da bu biraz (0.3) anlaması zor geldi/*I also found this a bit difficult to understand*, extract 12a). They also check to see if there are other possible correct answers in the item (e.g., doğru cevap kabul edilmesi gerekir orda/*it should be accepted as a correct answer there*, extract 6a) due to, for example, the possibility of the students' different interpretations of an expression or the use of an ambiguous term in the distractors (e.g., freely criticize edebiliyolar gibi algılayabilir/s/he can perceive it as though they can *freely criticize*, extract 9a) as well as the availability of relevant information in the reading/listening text (e.g., extracts 4a, 6a, 11a) or the possibility of the distractors within the context given in a multiple/binary-choice cloze test (e.g., extracts 3a, 7a). Raising such points while making an evaluation of a test item shows that the reviewers are aware of the assumed and shared testing principles that concern the design of language tests in general within their institutional context. Thus, they can approach the items as a tester, becoming an important part of the test construction process in their institution.

Thirdly, throughout the IR interactions in the phase of PI, the reviewers display an orientation to their institutional roles as teachers by, for instance, directing a question about the instructional events or students' L2 knowledge to the other reviewer, not the tester in the preliminary inquiry since as teachers, they have an epistemic access to the domain of teaching and the abilities and knowledge of their students. Similarly, as illustrated in Table 13.10, the reviewers happen to frame their claims or accounts from the students', teachers' and teaching perspectives. This implies the teachers' epistemic right, responsibility and obligation (Heritage, 2013) to review the item with consideration of the context of learning and the needs of students and thus, to create an interplay between teaching and testing. In addition, as illustrated in extract 11a, a reviewer uses a modified repeat (Stivers, 2005) of an assertion that concerns the students' knowledge of a word (i.e., *bilir* (.) *di mi?/s/he would know it, wouldn't she?*). This indexes her epistemic right to confirm an issue about the students' L2 knowledge, asserting her epistemic authority and control as the teacher with access to the domain of teaching and the target students' L2 abilities and knowledge (Heritage & Raymond, 2005; Stivers, 2005). Moreover, the reviewers' shared role as teachers with equal epistemic access to the domain of teaching is manifested in their use of an inclusive language for a teaching practice as in extract 12a (i.e., *ambitious öğretmen miydik/were we going to teach "ambitious"?*). Overall, those interactional resources indicate that the reviewers have a specific role and responsibility during the phase of PI in the IR interactions.

Despite an orientation to the institutional role distribution of the participants, they happen to share the responsibility for problematizing the item and for their final decision in such task-oriented IR interactions where the product is shared, affecting everybody in the group. Whether it is the reviewer or the tester, the participants all get engaged in the task of problematizing the item and follow a process of collaborative inquiry, actively and critically thinking about the item and making a contribution. This indicates that the participants show ownership and feel responsible for the test, the outcome of which is important for all of them. The involvement of all the participants in the process of problematizing the item is especially the case when the identified problem concerns an issue which they have equal epistemic access to. For instance, in extract 4a, since the problem is related to the options deriving from a reading text, which is an accessible source of information for all the participants, they collaboratively investigate the matter at hand by orienting to the reading text in an attempt to prove the problem during the IR interaction. They do not just leave it to

the problematizer of the item to provide an account but they themselves seek evidence, give an account from different perspectives and present their views with regard to the acceptability of the problem, which all play an important role in the final decision about the item.

The participants' equal active engagement in the phase of PI is most evident in their gaze behaviour. It is seen that the participants continuously look at their exam papers only sometimes directing their gazes at each other for a short time. Although a disengaged gaze would make it sanctionable in an ordinary conversation (Goodwin, 1981; Rossano, 2013), in such a task-oriented interaction, it does not indicate disengagement; on the contrary, it signals that the participants are all engaged in the task and actively think about the matter at hand. With their gazes on their papers, they, for instance, read the relevant part in the text, scan through it in search of evidence, shift their gazes between the text and the item, turn over the page, translate the target sentence in the text into Turkish, underline the possible correct answers or the relevant sentences in the text, check their notes on the note-taking sheet and orient to the tapescript in a listening test, point at the words with a pencil, circle the key words, and produce a thinking face. All these verbal and non-verbal actions while the participants gaze at their papers signal their role as active agents during the phase of PI. Even when a participant seems to be relatively more passive in the talk-in-interaction (e.g., AV in extract 5a), her orientation to the test material and her sustained gaze at her paper throughout the IR interaction show her engagement in the task.

The participants' shared role in the phase of PI is further evident in the way the reviewers announce the problem in the item as they leave room for collaborative thought and inquiry on the raised issue (extracts 2a, 4a, 5a, 6a, 7a, 9a, 10a, 11a, 12a). The problematizers do not state the problem assertively but they, for instance, present it through a direct or an embedded question for everyone in the group to think about as in extracts 4a, 6a and 7a, and/or initiate a pre-expansion sequence where they all preliminarily search for evidence for the problem or reconsider the item. As a result, the other participants also take an initiative to provide an account and state the problem by means of such interactional resources as candidate understanding of the problem, choral co-production, anticipatory completion, collaborative turn completions, candidate evaluation of the item, (re)formulations, and modified repeats (See Table 11). Finally, the participants' use of an inclusive language while projecting the phase of SC as the relevant next action (extracts 8a, 12a), or while announcing

the problem (extracts 6a, 12a) is another indication of their shared roles in the IR interaction. It demonstrates that all the co-participants have the epistemic right, responsibility and obligation (Heritage, 2013) to contribute to the item change process and to collaboratively decide on the problem as well as the changes to be made in the item.

#### 4.2. Phase 2. Suggesting a Change (SC)

Extract 1b presents the phase of SC in the reviewing interaction on the item problematized in extract 1a. It is to be remembered that in this extract, the two teachers BA and DE, who teach at B2 level, review a multiple-choice reading comprehension item in a level achievement test of B2 groups and provide their feedback to the tester EK. The reading comprehension item has a negative root (i.e., *Which of the following is NOT one of the advantages of the Kalenjini?*) and has four options. In the first segment of the extract (i.e., extract 1a, see Section 4.1), BA problematized option “b”, which is the correct answer, as she believes that it is too obvious and easy to spot for their group of students at B2 level because it is not at all related to the content of the reading text and can easily be regarded *not* as an advantage of the Kalenjini. Thus, in the following segment of the extract (i.e., extract 1b), having already announced and agreed with the problem, the IR group revises option “b” through a process of suggestion-making for change and improvement in the problematic option. This extract was also presented at a data session. The analysis provided here is, thus, supported by the analysis carried out at this data session.

### Extract 1b. Special Diet, Segment b

- 1 BA: hani (.) onun yerine ne [diyebiliriz acaba.  
*well I wonder what we can say instead of that*  
 + all look at their exam papers
- 2 DE: [professionals:  
 + EK moves her pencil along the option  
 + shifts gaze to BA
- 3 (1.4)  
 4 olmaz (.) gene  
*still not*
- 5 BA: ↑gene aynı şe:y.  
*it is still the same thing*  
 + EK turns gaze to the reading text, then back to the item
- 6 ((DE turns back to her paper and all sustain gaze at  
 their papers for 1.6s))

7 EK: şey diye↓lim (0.2) um::  
**let's say this**  
+ shifts gaze to the reading text  
+ DE shifts gaze to EK

8 (0.8)

9 DE: ↑older people=

10 EK: =following a special [diet- ]  
+ reads aloud the first part of the option

11 DE: [mese↑la]  
**for example**  
+ shifts gaze to BA  
+ BA shifts gaze to DE, pointing at her

12 BA: °uh-huh°  
+ BA and DE both shift gaze to EK  
+ EK produces thinking face

13 (0.8)

14 DE: ↑older people.

15 (0.7)

16 hani old-  
**you know**

17 EK: °nasıl?°  
**what?**  
+ shifts gaze to DE

18 DE: older people of the (0.3) tribe.  
+ EK turns gaze back to her paper

19 (0.9)

20 EK: designed by?  
+ shifts gaze to DE

21 (0.2)

22 DE: the [older- (.) old-] (.)  
+ BA turns back to her paper and nods head

23 EK: [the elderly ]

24 DE: [THE ELDERLY]  
+ EK turns gaze back to her paper  
+ shifts gaze to BA, EK and then back to her paper respectively

25 BA: [THE ELDERLY]  
+ nods head

26 (0.3)

27 °following a special diet designed by the <elderly>°  
+ reads aloud the suggested option  
+ EK shifts gaze to BA

28 ((DE crosses out the part of the option that is to be changed))

- 29       yani ordaki ex↑perts (.) ba↑na  
           **well to me "experts" there**  
                   + shifts gaze to DE  
                   + DE shifts gaze to BA  
                   + EK turns gaze back to her paper
- 30       (0.2)
- 31       böyle (.) [er: çok-  
           **is like too-**
- 32   DE:               [>anladım daha< professio[nal geliyo.=  
                       **I see, it sounds more professional**  
                                   + turns gaze back to her paper
- 33   EK:               [şey diyelim?=  
                       **let's say this**  
                       + shifts gaze to DE, BA and then  
                       back to her paper
- 34   BA: =şey geldi.  
           **sounds thing**  
           + turns gaze to EK
- 35   EK: şey [olur mu?  
           **would this thing be alright?**
- 36   BA:       [ne yap[alım  
           **what shall we do?**
- 37   DE:               [AMA Bİ DAKKA olderlar (0.3) yapar mı  
 38       acaba=  
           **but wait a minute, I wonder if the old would do it.**  
                                   + BA turns gaze back to her paper
- 39       =↓yok.  
           **no**
- 40       (0.2)
- 41   EK: following a special diet consisting of  
           + reads aloud her suggestion for the option  
           + DE shifts gaze to EK, then back to her paper
- 42       (0.9)
- 43       ↑meat  
           + shifts gaze to BA  
           + BA shifts gaze to EK
- 44       (1.0)
- 45   BA: [ °huh °
- 46   DE: [e za- ama o-  
           **but that's**  
           + EK turns gaze to DE, then back to her paper
- 47       (0.8)
- 48       huh
- 49       + BA shifts gaze to DE



50 EK: böyle bi bilgi yok.  
*there is no such information*  
 + BA turns gaze back to her paper  
 + DE directs gaze towards the reading text

51 DE: yok mu?  
*isn't there?*

52 (0.2)

53 BA: tabi yok. (.)  
*of course not*

54 EK: [çok et]  
*too much meat*

55 DE: [tamam ]  
*okay*

56 EK: yiyolar [o yüzden-]  
*they eat because of that*

57 DE: [following] a spe[cial [↑DIET ]  
 + reads aloud the new option + crosses out the changed part of the option

58 BA: [cial [↑DIET ]

59 EK: [°diet°]

60 (1.2)

61 BA: <consisting>  
 + EK writes down the new option

62 ((DE writes down the new option and BA produces thinking face))

63 DE: of:  
 + reads aloud the new option while writing it down

64 (0.5)

65 meat.

Extract 1b begins with BA's embedded question which asks for a suggestion to change option "b" as all the participants are looking at their exam papers. Her question acts as a preface which projects the upcoming suggestion and decision making process for the problematic option and initiates a new sequence. This at the same time signals readiness to start item modification. Just prior to the completion of BA's question, DE immediately provides the alternative "professio↑nals:" for "experts" in the option with a pitch rise in word-final position and an emphasis on the final syllable of the word while EK moves her pencil along the option. Through the end of her TCU, DE also shifts her gaze to BA but a gap of silence ensues with no uptake from the other participants as they all think about the offered alternative. Then in line 4, DE herself disagrees with her suggestion and expresses her dispreference through a self-cancellation in the form of a self-initiated self-repair. In this

respect, she backs down and withdraws her suggestion. BA agrees with DE's stance and shows alignment by starting her turn with the last word of DE's turn (i.e., *↑gene/still*) with an initial pitch rise and asserting that it is not different from the original option. At the same time, EK turns her gaze to the reading text and then back to the item in search of an alternative for the option. This is followed by a lapse of 1.6 seconds when DE turns back to her paper and they all sustain their gazes at their papers, thinking about new possibilities to improve the option.

Then in line 7, EK takes over and draws the other participants' attention to an upcoming suggestion by applying a preface with an inclusive expression of suggestion (*şey diye↓lim/let's say this*) accompanied by her shift of gaze to the reading text. As a result, DE directs her gaze towards EK. However, EK does not give her suggestion yet and uses an elongated hesitation marker followed by a gap of brief silence as she is thinking about a possible alternative for the option. Before she comes up with an idea, DE takes a turn in line 9 and provides the second suggestion "*↑older people*" to replace "*experts*" in the option with a rising intonation in word-initial position. With no time lapse, in disalignment with DE's suggestion, EK continues to further think about an alternative by reading aloud the first part of the option in line 10 and stopping before the trouble source "*experts*" in the option and then producing a thinking face to complete the option with an alternative. The other co-participants DE and BA do not orient to EK but maintain mutual gaze in line 11. BA acknowledges DE's suggestion (i.e., "*older people*") by pointing at DE and uttering an acknowledgement token quietly. Only then BA and DE shift their gazes to EK and after a gap of brief silence, DE reminds EK of her suggestion in line 14 with a rising intonation followed by her initiation of an explanation for her suggestion.

Before DE completes her account, EK cuts off her utterance in line 17 and shifting her gaze to DE, she initiates a repair action by quietly asking DE to reproduce her suggestion. In response to the other-initiated repair, DE produces the repair solution in line 18 by repeating her suggestion and then extending it with "*of the tribe*" to make it more meaningful as she is still in the process of formulating her suggestion. Just prior to the completion point of DE's unit of talk, EK turns her gaze back to her paper. In line 20, shifting her gaze to DE again, she initiates another repair action with respect to the position of DE's suggestion in the original option by reading aloud "*designed by?*" in the option with an interrogative

intonation and constructing a Designedly Incomplete Utterance (DIU) (Koshik, 2002) to elicit DE's repetition of her suggestion and thus, to get her to place it meaningfully in the option as a replacement of the trouble source (i.e., *experts*). Following a brief gap of silence, just as DE attempts to restate her suggestion in alignment with EK's DIU, EK overlaps and delivers it herself but with a reformulation of DE's suggestion as she transforms the already offered suggestion "older people" into its rephrase "the elderly" in line 23. Meanwhile, DE completes EK's DIU with "the older" but then starts over by engaging in a self-initiated self-repair although she cannot complete it due to EK's overlapping turn. This demonstrates that DE is still in the process of formulating her suggestion.

BA acknowledges DE once again by turning back to her paper and nodding her head. However, EK's "the elderly" is a more preferred response for both BA and DE as they engage in a choral co-production (Lerner, 2002) by simultaneously repeating and co-producing "THE ELDERLY" at a higher volume in lines 24 and 25. Thus, they display agreement with the reformulated suggestion. At this point, EK turns her gaze back to her paper and during their choral co-production, DE first orients to the other co-participants and then back to her paper while BA embodies her turn with a head nod, which further signals her agreement with EK's "the elderly". In line 27, BA silently reads aloud the whole option by inserting the newly suggested word into it, checking and reviewing it and then DE crosses out the part of the option that is to be changed.

In lines 29 and 31, BA provides further clarification of the problem she has previously stated with regard to the item by once again referring to "*experts*" in the original option as the trouble source with an emphasis and a rising intonation on it and engaging in mutual gaze with DE. Just as BA uses an elongated hesitation marker in search of an explanation, DE employs anticipatory completion (Lerner, 1996a) in line 32 before BA completes her turn. In doing this, DE first claims understanding and then demonstrates it by explaining that "*experts*" in the option sounds more professional with a shift of gaze back to her paper. Overlapping DE's turn at a turn-final position, EK projects another upcoming suggestion by applying a preface with an inclusive expression of suggestion and a slight rising intonation in line 33 and pre-emptively requesting the other participants' confirmation for her upcoming suggestion in line 35. At this point, EK orients to the other co-participants with a shift of gaze to them but then turns to her paper in search of a suggestion. BA completes her

prior turn in line 34 in alignment with DE's anticipatory completion and turns her gaze to EK, asking her for the suggestion she has projected in the meanwhile (line 36). However, before EK provides her suggestion, DE displays disalignment with a *but*-prefaced turn uttered at a higher volume and does not orient to EK's projected suggestion. Instead, DE plays the devil's advocate in line 37 and with her gaze on the reading text, she challenges the suggestion she has previously offered by asking whether the old people are also said to design a special diet in the reading text. Just prior to the completion of her turn, BA turns her gaze back to her paper. DE then immediately delivers the answer to her question herself through a negative response token (i.e., ↓yok/no) uttered with a falling intonation. Thus, she engages in a self-cancellation by withdrawing her claim, which presented a challenge to her suggestion.

Then EK takes over and reads aloud the option by replacing the second component of the option (i.e., *designed by experts*) with her new suggestion "consisting of ↑meat" with an initial rise in the pitch of the word "meat". She produces "↑meat" following a lapse of 0.9 seconds, which demonstrates her online thinking process about her suggestion. During her turn-in-progress, DE shifts her gaze to EK and then back to her paper and EK maintains mutual gaze with BA just as she utters the new word "↑meat". After a gap of silence, BA responds with a quietly uttered acknowledgement token while DE hesitates about EK's suggestion and challenges it in line 46 as evident in her use of "ama/*but*" as a sign of disagreement and intra-turn cut-offs. At this point, EK directs her gaze towards DE and then back to her paper. Following a gap of silence, BA also shifts her gaze to DE. Then in line 50, initiating a non-minimal post-expansion sequence, EK provides an account for her suggestion as a counter-challenge to DE's initiation of resistance to EK's suggestion by claiming that there is no such information in the text. Simultaneously with EK's turn, BA turns her gaze back to her paper and DE directs her gaze towards the reading text, checking if the new suggestion would be right and considered as an advantage within the given context in the reading text because the item asks for an option that is not an advantage mentioned in the reading text. EK's assertion is met with DE's confirmation request in line 51 with regard to the availability of the newly suggested information in the reading text. DE receives a confirming response from BA in line 53 which is marked with the expression of certainty "tabi/*of course*" (Stivers, 2011).

In lines 54 and 56, EK provides further account for her suggestion and overlapping her at a turn-initial position, DE this time shows agreement with the suggestion via the minimal expansion acceptance token “tamam/okay” in line 55 and cuts off EK’s utterance in line 57 when she starts reading aloud the new option. BA then engages in choral co-production (Lerner, 2002) with the current speaker DE. BA co-produces the final syllable of “special” and “↑DIET” at the same time and in the same manner as DE since they both articulate “↑DIET” at a higher volume and with a pitch rise in word-initial position. DE’s reading aloud is embodied with her crossing out the changed part of the option on her exam paper. EK also joins their choral co-production in line 59 by uttering “°diet°” quietly at the same time with them. Once they read out the first component of the original option, a gap of silence ensues and between lines 61 and 65, BA and DE together replace and complete the rest of the option with the new suggestion by reading it aloud through collaborative turn completion while EK and DE write it down on their exam papers. Eventually, the participants reach an agreement on the suggestion offered by EK and a change is made in the problematic option.

Extract 1b demonstrates the suggestion and decision making process to change and improve a multiple-choice reading comprehension item, which was problematized previously in the first phase of the IR interaction. In order to make the correct option “b” less obvious to select as an advantage not mentioned in the reading text, the participants offer some suggestions until they decide on one of them. It is seen that one suggestion paves the way for another suggestion in the talk-in-interaction as the participants are constantly in search of a better alternative. In this particular extract, four different suggestions are offered to replace the trouble source “experts” in the option, which are (1) *professionals*, (2) *older people*, (3) *the elderly*, and (4) *consisting of meat* and the group decides on the final one (i.e., *consisting of meat*). The decided alternative is different from the others in that it is not a person designing the special diet stated in the option but it is about the content of the diet and thus, changes the trajectory of the problematic option, making it more acceptable by the participants.

A collaborative process is followed via some interactional actions while seeking a change for the option and the suggestions are not simply listed one by one as distinct entities. First of all, it is seen that a suggestion that is offered by one participant can be reformulated and

redesigned by another participant through the talk-in-interaction. In this extract, the suggestion “*the elderly*” emerges as the reformulation of the previous suggestion “*older people*” through a repair action and the use of a DIU, which shows that the suggestions are collaboratively shaped in and through talk. Likewise, EK’s repair action enables DE to redesign her prior suggestion “*older people*” as “*older people of the tribe*” in an attempt to make it more meaningful and understandable by the other participants.

Secondly, the participants check and challenge a suggestion before orienting to the upcoming ones and making a decision. In this extract, the challenging is especially done by DE, who has the role of playing the devil’s advocate and questioning not only the others’ but also her own suggestions in an attempt to be on the ‘safe’ side and in order not to make changes that would make the option even more problematic. For example, she challenges her suggestion “*older people*” as well as EK’s suggestion “*consisting of meat*” with respect to their availability in the text so that the option does not end up being expressed as an advantage in the text and thus, being an incorrect option since the item asks for an advantage not mentioned in the text. Similarly, DE and BA both challenge the first suggestion “*professionals*”, which leads to the realization that it retains the violated testing principle that the option is too obvious for the students at B2 level to select.

As a result of the act of challenging a suggestion, the participants either eliminate the offered alternative by backing down and withdrawing it as in the case of “*professionals*” (i.e., self-cancellation) or keep it as an alternative to consider. In the latter case, the participants either withdraw the claim that has questioned the suggestion after checking and disproving it as DE does in the case of “*older people*” or make a counter-challenge in a non-minimal post-expansion sequence by conjointly providing an account for the suggestion as BA and EK do in the case of “*consisting of meat*” and then reaching an agreement. Thus, the act of challenging a suggestion results in the withdrawal of the claim made against the suggestion and account-giving for the suggestion, which appear as the other interactional actions that contribute to the change process in the item in the phase of SC. In this respect, the act of challenging a suggestion by both the provider of the suggestion and the other participants in response to the suggestion can be considered to be an important contribution to the validation process since it prompts a critical and careful thought on the suggestion in and through talk (Fulcher & Davidson, 2007). As a result, such a collaborative way of understanding the suggestions leads to the final decision with respect to the change to be

made in the problematic item “in open discussion, through a process of dialogue and disagreement” (Fulcher & Davidson, 2007, p.21).

Thirdly, this extract illustrates that even when an agreement is displayed with a suggestion, the participants seek other alternatives and then decide on a change. For instance, although a preference is shown for EK’s “*the elderly*”, EK still continues to search for an alternative and eventually, formulates a different suggestion (i.e., *consisting of meat*) that is met with the other’s uptake and decided to be inserted into the option as a replacement of the trouble source. In relation to this, the participants display agreement with a suggestion collaboratively by employing choral co-production (Lerner, 2002), reading aloud the option with the suggestion inserted into it, and writing down the new option on their exam papers through collaborative turn completions, which at the same time enable a final check for the suggestion and ensure shared understanding regarding the change made in the item.

Fourthly, the decided suggestion that is offered by EK is downgraded since it is directed to the other participants as a question in the form of a pre-emptive confirmation request in the preface (line 35). In this way, EK does not implicate epistemic priority as the first provider of a new suggestion while the reviewer DE has also already offered a suggestion awaiting decision (Heritage & Raymond, 2005) but invites collaborative inquiry into her suggestion, sharing the responsibility for the decision. Since the participants all have equal epistemic access to the reading text, such an approach inviting collaborative inquiry would allow for a careful thought on the suggested changes and thus, a ‘healthier’ decision-making. Eventually, following the offer of EK’s suggestion, a gap of silence ensues and then the group starts to discuss her suggestion. On the other hand, EK and BA’s display of certainty with their accounts and claims in support of the suggestion as evident in their declarative statements and BA’s marked confirmation through the use of “*tabi/of course*” (Stivers, 2011) lead to the acceptance of EK’s suggestion by the group. Finally, another interactional resource emerging in the phase of SC is the participants’ use of a preface which projects the announcement of an upcoming suggestion. In this way, they draw the other participants’ attention to the suggestion as they shift their gazes to the participant who has used a preface. BA does this through an embedded question as the sequence-initiating action for the phase of SC to elicit suggestions for the problematic option while EK applies a preface as a projection of a new suggestion that she is in the process of formulating.

This extract also shows an orientation to the reviewer roles and responsibilities in the IR interaction. To begin with, at one point in the interaction as the suggestions for change are offered, the reviewer BA who identified the problem in the item in the phase of PI provides further clarification of the problem, reminding the group of the matter at hand and ensuring shared understanding with respect to why they are to change the option. In this way, as the reviewer, she monitors the talk-in-interaction, making sure that the participants know what to correct in the item and design their suggestions accordingly. In addition, throughout the IR interaction, DE plays the devil's advocate by challenging her own as well as the others' suggestions and requesting confirmation, which orients to her epistemic right, responsibility and obligation (Heritage, 2013) to carefully evaluate each suggestion as the reviewer of the test and confirm it only when there is solid evidence for it.

With respect to the tester roles, as the tester who is accountable for the test, EK considers the change suggested by DE (i.e., *older people*) carefully by constructing a DIU (Koshik, 2002) to get DE to place it meaningfully in the option as a replacement of the trouble source (i.e., *experts*) and checking the appropriacy of the suggestion within the context given in the option. Throughout the interaction, she is also in search of a better alternative for the problematic option since she rephrases DE's "*older people*" as "*the elderly*" which is met with a display of preference by the other participants and since she then formulates a new alternative that is accepted by the group as the change to be incorporated into the option.

Although this extract displays an orientation to the institutional role distribution of the participants, they happen to share the responsibility for suggesting a change in the item and for their final decision in such a task-oriented IR interaction where the product is shared. This is most evident in all the participants' constant search and online formulation of an alternative to overcome the problem in the item. The tester EK, for instance, makes her online thinking process clearer by means of her elongated hesitation markers following a preface that projects an upcoming suggestion, gaps of silences while formulating her suggestion, reading aloud the first component of the option and then stopping before the trouble source as well as her production of a thinking face. As she is fully occupied searching for a suggestion, she does not even orient to the alternatives offered by DE in the first place. Similarly, DE also shows her active thinking process about a suggestion through the gaps of silences that follow her suggestions just before making a comment on them, intra-turn cut-offs and pauses while formulating a suggestion, her redesign of her suggestion



as when “*older people*” becomes “*older people of the tribe*”, and self-cancellation of her suggestion as well as her claim that challenges her suggestion. Thus, she continues to think about the suggestion even after producing it without leaving it to the other participants to evaluate it since they all have equal access to the reading text and can evaluate the suggestions based on the information in the text.

The engagement of all the co-participants in the task of suggesting a change is also evident in their gaze behaviour. It is seen that throughout the interaction, the participants continuously look at their exam papers only sometimes directing their gazes at each other for a short time but this does not indicate disengagement; on the contrary, it signals that the participants actively think about an alternative for the problematic option in such a competing activity (Goodwin, 1981; Rossano, 2013). Their constant shift of gaze to the reading text as well as the item and non-verbal actions such as moving the pencil along the option, crossing out the part of the option to be changed and marking the change on the exam paper all indicate the participants’ right and responsibility to seek a change to improve the item and decide on it after carefully evaluating it. They also demonstrate the participants’ ownership of and shared responsibility for the test, the outcome of which is important for all of them. Finally, the participants’ use of an inclusive language while projecting a suggestion in the preface once again makes a reference to their shared responsibility for the change process in the item.

Extract 2b presents the phase of SC in the reviewing interaction on a grammar item in a tense passage which was problematized in extract 2a. It is to be remembered that in this extract, the two teachers YE and CI teaching at B1 level do the reviewing of the item and provide their feedback to the tester SA. In the first segment of the extract (i.e., extract 2a, see Section 4.1), YE problematized the verb “*cover*” given in one of the blanks in the tense passage to be transformed into its past tense form, claiming that it was not studied in class. Thus, in the following segment of the extract (i.e., extract 2b), having already announced and agreed with the problem, the IR group finds an alternative for the problematic verb through a process of suggestion and decision making for change and improvement in the item. This extract was also presented at a data session. The analysis provided here is, thus, supported by the analysis carried out at this data session.

## Extract 2b. Cover, Segment b

- 1 YE: daha iyi bi alternatif <olabilir  
*there can be a better alternative*  
 + looks at her exam paper  
 + CI shifts gaze to YE  
 + SA looks at YE
- 2 ama> change falan olursa  
*but if something like "change" is possible*  
 + directs gaze towards CI  
 + SA turns gaze to her paper
- 3 [belki] [change]=  
*maybe*
- 4 CI: [hu:h ]
- 5 SA: [hu:h ]
- 6 YE: =geçmiştir [diye düşündüm emin değilim ama ]  
*appeared I thought but I'm not sure*
- 7 CI: [tabi daha iyi bilirler change'i]  
*of course they would know "change" better*  
 + turns gaze to her exam paper by nodding head  
 + nods head
- 8 SA: change [diyelim mi? o zaman ona]=  
*then shall we make it "change"*  
 + YE shifts gaze to SA's exam paper
- 9 CI: [evet change'i daha iyi ] [bilirler]  
*yes they would know "change" better*
- 10 YE: =[evet ]  
*yes*
- 11 SA: <sup>1#</sup>tamam (0.2) <yani hani> verb ikisini bilemezler (.)  
 + CI writes the new verb  
 "change" on her paper
- 12 [mi acaba diye<sup>#1</sup> diyosunuz]  
*okay well you know do you say so because they may not know its verb two*  
 #1 (lines 11-12) SA makes the change on her exam paper by writing down the new verb "change"
- 13 YE: [evet aynen öyle ] [hani-  
*yes that's right well-*
- 14 CI: [(°bunu: er°)  
*this*
- 15 SA: [tamam hocam  
*okay miss*
- 16 YE: benim ↑derdim o [change]  
*that's my concern, that "change" is*  
 + SA shifts gaze to CI and then to YE
- 17 SA: [tamam ]  
*okay*

18 YE: olmasa bile belki başka-  
*even if not possible, maybe something else*  
 + turns gaze back to her paper

Extract 2b begins with the reviewer YE's preface for the phase of SC as she is looking at her exam paper while the tester SA's gaze is at her. By articulating that a better alternative can be found for the verb "*cover*", she projects the change process for the improvement of the item as the relevant next action in the IR interaction and initiates a new sequence where suggestions for change are expected. At this point, CI also shifts her gaze to YE. Following her preface, YE herself quickly offers an alternative (i.e., *change*) for the problematic verb "*cover*" with an emphasis on it by directing her gaze towards CI and maintaining mutual gaze with her. Just after YE announces her suggestion, SA turns her gaze to her exam paper. YE's suggestion is immediately received with an elongated minimal acknowledgement token (i.e., *hu:h*) from both CI and SA before she completes her turn. In line 6, without any time lapse, YE continues with her turn and completes the turn constructional unit which she initiated in line 3, explaining that the verb she suggests could have been studied in class. She downgrades her suggestion by using the hedge "*belki/maybe*" and further expressing uncertainty about whether the offered verb "*change*" was studied in class. In this way, she also checks her account for the suggestion, eventually receiving a confirming response from CI in line 7, which is at the same time embodied with CI's head nod while shifting her gaze to her exam paper.

Regarding her confirming response, CI displays alignment and agreement with YE's account in an overlapping turn by providing a parallel assessment (Heritage, 2011) of the suggestion (i.e., *daha iyi bilirler/they would know better*) but with a greater certainty through the use of the marked confirmation "*tabi/of course*" (Stivers, 2011). Thus, CI conveys a stronger evaluative stance towards the suggestion than YE. In addition, while YE presents her assessment of the suggestion from the teaching perspective based on evidence from a shared past instructional event (i.e., *geçmiştir/appeared*) (Can Daşkın & Hatipoğlu, 2019; You, 2015), CI frames hers from the students' perspective in more general terms with the use of present tense (i.e., *bilirler/they would know*).

Following CI's affiliative assessment of and account for the suggestion, SA directly displays her acceptance of the suggested verb "*change*" and her readiness to change the item through

her offer to replace the problematic verb “*cover*” with the new verb “*change*” with an inclusive expression of suggestion (*diyelim mi?/shall we say*) in the form of a confirmation request (line 8). Just as SA launches her turn, YE shifts her gaze to SA’s exam paper and then CI overlaps in line 9 with the compliance token “*evet/yes*” in turn-initial position and repeats her prior assessment of the suggestion. This further acknowledges and confirms the appropriacy of the suggested verb “*change*” in terms of the testing principle that the students should know the given verbs in a tense passage. Then immediately following SA’s turn, YE provides a confirming response to her offer to change “*cover*” with “*change*” in line 10 via a minimal compliance token. This is, in turn, met with SA’s agreement in line 11 through her use of a minimal expansion acceptance token (*tamam/okay*). They eventually decide on the first offered verb as a replacement of “*cover*” without further seeking other alternatives.

Following this suggestion and agreement sequence, which is closed with the decision to change the item with the suggested verb, SA initiates a non-minimal post-expansion sequence in line 11 just as she notes the new verb on her copy of the exam paper. In this post-expansion sequence, SA offers a candidate understanding (Antaki, 2012; Heritage, 1984b) of what the reviewers have previously stated regarding the problem in the item and seeks for their confirmation. In this way, she further checks why they have made such a change in the item and ensures shared understanding in the group. During SA’s turn-in-progress, CI also marks the change on her exam paper by writing down the new verb “*change*”. Overlapping SA’s turn at a turn-final position, YE confirms SA’s candidate understanding as correct in line 13 and then in lines 16 and 18, she further elaborates on the issue by reasserting that her concern is with the verb in the item and adding that it could be substituted with any other verb even if not the verb “*change*”. In line 14, CI also attempts to say something with a softer voice but does not complete it. YE’s confirmation of SA’s candidate understanding is received with SA’s use of a minimal acceptance token followed by an address term (i.e., *tamam hocam/okay miss*), which shows SA’s agreement and affiliation with YE’s stance. With the repetition of the same acceptance token in line 17 in overlap with YE’s ongoing clarification, SA then proposes sequence closure.

Through a process of suggestion and decision making, extract 2b demonstrates how the participants interactionally deal with the problem identified in the first phase of the IR

interaction with respect to the verb “cover” given in a tense passage. In this extract, the group focuses on one particular suggestion (i.e., the verb “change”) as offered by the reviewer YE and the interaction revolves around its evaluation, according to which a decision is made in the end. In this respect, this extract shows the important role of collaboratively providing account for and making assessments of the suggested changes in the decision-making process to improve a problematic item. To illustrate, before deciding on the suggested verb “change”, it is evaluated from different perspectives and in each case of evaluation and account giving, an orientation is shown to the violated testing principle “test what you teach” and thus, to the participants’ situated understanding of how a test should be. YE, for instance, presents her view of the suggestion from the teaching perspective by making reference to a shared past instructional event (Can Daşkın & Hatipoğlu, 2019; You, 2015) where the verb “change” is assumed to have been studied while CI considers it from the students’ perspective by referring to their present state of knowledge concerning the verb “change”. In any case, the suggestion is received with a positive assessment from the reviewers in the form of a parallel assessment (Heritage, 2011) and is not challenged.

YE and CI’s parallel assessment of the suggestion with different evaluative stances has some interactional implications. For instance, during her evaluation of her suggestion, YE downgrades her epistemic stance by showing less certainty with respect to the teaching of the verb she has offered (i.e., *belki change geçmiştir diye düşündüm emin değilim ama/I thought maybe “change” appeared but I’m not sure*) whereas CI asserts greater epistemic authority regarding the students’ knowledge of the verb by presenting her claim (i.e., *tabi daha iyi bilirler/of course they would know better*) as assumed knowledge through the use of the stance marker “*tabi/of course*” (Stivers, 2011) and the present tense, which signal certainty. This could be because YE is both the ‘owner’ and the first assessor of the suggestion. YE might be downgrading her epistemic claim in order not to implicate epistemic priority as the first assessor or provider of the suggestion (Heritage & Raymond, 2005) and to be able to invite others’ contributions to the evaluation process of her suggestion, sharing the responsibility for the decision. Since both YE and CI have equal epistemic access to the referent situation (i.e., the teaching context), through her account check, YE might not want to block CI’s contribution by being too assertive about her suggestion but rather leave room for her assessment of the suggestion as well, which would allow for a careful thought on the

offered verb “*change*” and a ‘healthier’ decision-making. In addition, while YE’s downgrading of her assertion paves the way for further collaborative thought on the suggested verb, CI’s stronger evaluative stance that the students would certainly know the verb “*change*” as well as her repetition and strong confirmation of her account in line 9 lead to its immediate acceptance by SA. Upon CI’s evaluation, the appropriacy of the verb is assured in terms of the testing principle “test what you teach” and the verb is not further questioned, being directly substituted with the problematic one on the exam paper without a need to seek other alternative verbs.

As for the institutional role distribution of the participants in extract 2b, there are indications of reviewer/teacher and tester roles and responsibilities in this IR interaction. To begin with, the evaluation of the suggestion is considered as part of the reviewer/teacher responsibility and hence, the tester SA does not contribute to this process. This is also evident in YE’s gaze orientation to the other reviewer CI while providing the suggestion in the first place. This could be because the state of affair, that is the verbs the students would presumably know, falls into the teachers’ epistemic domain as they are best familiar with their students’ knowledge of the language they are learning and as part of their institutional roles, they have the right, responsibility and obligation to know and have access to their students’ epistemic domain of the target language (Heritage, 2013). This means that unlike SA, YE and CI have the right to evaluate the suggestion in terms of their students’ knowledge of L2 verbs based on what is taught to them in class so that the item can be co-constructed in line with the testing principle “test what you teach”. This finding is contrary to what was found in extract 1b. In this extract, although there was an orientation to the participants’ institutional role distribution, not only the reviewers but also the tester was actively involved in making and evaluating a suggestion due to the nature of the problem and the violated testing principle. For instance, in extract 1b, when the correct option in a multiple-choice reading comprehension question was problematized for being too easy for the target group of students to select, all the participants oriented to the test material in search of an alternative to overcome the problem because the problem was directly related to the reading material in the test, which all the participants had equal epistemic access to. This was the reason why the tester also offered suggestions and evaluated them by providing evidence from the reading text.

As for the tester roles in extract 2b, SA's confirmation requests throughout the interaction orient to her role as the tester who is accountable for the test. She, first, solicits confirmation with respect to the change to be made in the item in line 8, which forms the basis for the final decision. In this way, as the tester, SA monitors the item change process by directing the trajectory of the interaction towards the final decision upon the reviewers' account for the suggested verb "*change*" and by closing the suggestion-making sequence without allowing the reviewers to seek other alternative verbs. However, at the same time, through her confirmation request, she enables the reviewers to overtly confirm the decision on the item change, soliciting mutual agreement with them before marking the change on her exam paper. Secondly, SA initiates a post-expansion sequence through a confirmation request for her candidate understanding (Antaki, 2012; Heritage, 1984b) of the problem announced in the first phase of the IR interaction. In this way, she elicits further elaboration on the issue (Stivers, 2013) and ensures shared understanding in the group. SA's correct understanding of the problem and of why such a change is made in the test is important for accountability issues in testing because in any case of criticism, as the tester of the institution, she will have to explain to the stakeholders why the item is designed in such a way.

Despite an orientation to the institutional role distribution of the participants, they happen to share the responsibility for the change process in the problematic item. For instance, although SA was relatively more passive in the suggestion-making sequence due to her relative lack of epistemic access to the teaching context, by initiating a post-expansion sequence where she offers a candidate understanding (Antaki, 2012; Heritage, 1984b) of the problem, seeks confirmation for it and displays a situated understanding of the norm for testing, she actively engages in the interaction when made relevant, making different contributions to the interactional process. Likewise, SA's offer to change the item with the suggested verb "*change*" in the form of a confirmation request leads to joint decision-making in the IR group. In addition, the participants' shared responsibility for the item change and their active engagement in the task is evident in SA's use of an inclusive expression while making the final decision and in both SA and CI's non-verbal actions such as marking the change on their exam papers and shifting their gaze to their papers as they are thinking about the offered suggestion (Goodwin, 1981; Rossano, 2013). In short, this extract indicates that all the co-participants in the IR interaction are actively processing as they are suggesting and deciding on a change to improve the item and they somehow have a contribution to make to this change process.

Extract 3b presents the phase of SC in the reviewing interaction on a grammar item in a multiple-choice cloze test which was problematized in extract 3a. It is to be remembered that in this extract, the two teachers BA and DE teaching at B1 plus level do the reviewing of the item and provide their feedback to the tester ME. In the first segment of the extract (i.e., extract 3a, see Section 4.1), the reviewers problematized the item, which tested the students' knowledge of the conjunctions they had studied in class, claiming that there was more than one possible correct answer. Thus, in the following segment of the extract (i.e., extract 3b), having already announced and agreed with the problem, the IR group finds an alternative for the other possible correct answer “*therefore*” besides the expected answer “*furthermore*” through a suggestion and decision making process. This extract was also presented at a data session. The analysis provided here is, thus, supported by the analysis carried out at this data session.

### Extract 3b. Therefore, Segment b

- 1 DE: ↑WHEREAS diyelim therefore [yerine]  
**let's say "whereas" instead of "therefore"**  
 + all look at their exam papers  
 + has hand over mouth  
 + points at “therefore” in the options with a pen in her hand
- 2 BA: [evet ]  
**yes**
- 3 (0.2)
- 4 [whereas: (.) whereas ] diyelim.  
**let's say**  
 + shifts gaze to ME and then back to her paper
- 5 ME: [o↑lur niye olmasın tabi.]  
**okay why not, sure**
- 6 DE: o zaman furthermore (0.3) tam şey-  
**then "furthermore" is the exact thing**  
 + ME crosses out “therefore” on her paper and then replaces it with “whereas”  
 + DE and BA cross out “therefore” on their papers
- 7 ((BA and DE write “whereas” on their papers for  
 1.4s while ME is looking at her paper))

As all the participants are looking at their exam papers with DE's hand over her mouth, extract 3b begins with DE's announcement of her suggestion “*whereas*” as an alternative for the distractor “*therefore*” which was found to be also possible as a correct answer. DE uses an inclusive expression of suggestion and thus, with her suggestion, DE initiates the change process to overcome the identified problem in the grammar item. She makes her initiation of



a new sequence clearer by producing her offer of the conjunction “*whereas*” at a higher volume with an initial rise in its pitch and an emphasis on its final syllable. During her turn-in-progress, DE’s articulation of the problematic option “*therefore*” is also embodied as she points at this conjunction given in the options on her paper with a pen.

DE’s suggestion is met with an immediate responsive uptake by BA, who accepts the new alternative through a compliance token (*evet/yes*) in line 2 and then shifting her gaze to ME, thinks about the offered conjunction “*whereas*” by repeating it in line 4 with an elongation at word-final position. Simultaneously with BA’s turn, ME also provides a positive response to DE’s suggestion in line 5 through her use of the acceptance token “*o↑lur/okay*” with a rising intonation, her positive comment (*niye olmasın/why not*), and her expression of certainty via the marked confirmation “*tabi/sure*”. ME’s positive response enables BA to verify her acceptance of the suggestion as BA then offers to change the problematic option with “*whereas*” with an inclusive expression of suggestion (*whereas diyelim/let’s say “whereas”*) following a micro-pause in her turn-in-progress while shifting her gaze back to her paper. Following the group’s agreement and affiliation with the suggested conjunction “*whereas*”, the participants display their uptake of the suggestion by crossing out the problematic distractor “*therefore*” on their exam papers and replacing it with the newly offered conjunction “*whereas*”. This further indicates their acceptance of and decision on DE’s suggestion for change in the item. As they mark the change on their papers, DE also comments on their decision in line 6 by stating that “*furthermore*” is now ensured as the correct answer although she cuts off her utterance before it is complete. In this way, she shows that the problem in the item is now overcome.

Different from the other extracts, in extract 3b, the problem identified in the grammar item is dealt with in a more straightforward way. A suggestion is offered by the reviewer DE (i.e., *whereas*) and then directly accepted by the other participants as an alternative for the problematic distractor “*therefore*”. In other words, the suggestion is quickly followed by the decision to change “*therefore*” with “*whereas*” without further being subject to the act of challenging and account-giving. The participants also do not seek other alternative conjunctions once they adopt an affiliative stance towards “*whereas*”. This could be explained in two ways. First of all, the suggestion is offered with certainty as DE produces it at a higher volume with an initial rise in its pitch and an emphasis on its final syllable. This

demonstrates that DE came up with the idea after giving it a careful thought with her gaze at her paper and her hand over her mouth. Likewise, DE's final comment on the decision where she states that "*furthermore*" is now ensured as the only correct answer is another indicator of her confidence with respect to the change made in the item. This is unlike in extract 1a where the suggestions were offered quickly and some were formulated on spot during the turn-in-progress, as a result of which they were challenged by both the provider of the suggestion and the other participants.

Secondly, the participants might have quickly evaluated the newly offered conjunction non-verbally considering the context given in the cloze text with their gazes already at the item on their exam papers since they all have equal epistemic access to this text and the problem in the item concerns the possibility of one of the distractors (i.e., *therefore*) within the context given in this text. This is also evident in BA's repetition of the suggested conjunction "*whereas*" with an elongation at word-final position, which shows that she is thinking about the suggestion before the final decision. Thus, they might have accepted the suggestion after quickly evaluating "*whereas*" individually and finding out that it poses no risk at all when compared to the problematic distractor "*therefore*". Since they all display a stance of collaborative agreement with the appropriacy of the suggestion as a distractor, they do not explicitly provide an account but just decide on it.

As in the case of other extracts, extract 3b demonstrates that the participants share the responsibility and follow a collaborative decision-making process for the change in the problematic item. This is first evident in DE and BA's use of an inclusive language while offering to change "*therefore*" with "*whereas*" and in all the participants' marking the change separately on their exam papers by crossing out "*therefore*" and replacing it with "*whereas*", which all illustrate their shared responsibility for the decision and their ownership of the test as the shared product. In addition, the participants' gaze behaviour signals that they are all actively engaged in the task (Goodwin, 1981; Rossano, 2013) as they continuously look at their exam papers throughout the talk-in-interaction while the suggestion is offered or responded to. The participants' gaze behaviour further indicates their right, responsibility and obligation to think about the suggestion before displaying an uptake of it.



4 SA: ötü[rü< doğru olur ]  
*because of, it would be correct*

5 GE: [((clicks tongue)) o da olmuyo.]  
*it is not possible*

6 (1.2)

7 DU: don't do shopping dersin  
*you would say*

8 (1.7)

9 o za[man e ]lenmiş olur  
*then it would be eliminated*

10 SA: [o zaman]  
*then*

11 tek (.) doğru- tek er olumsuz şık o olur.  
*the only correct- it would be the only negative option*  
+ shifts gaze to DU and then back to her paper

12 (2.4)

13 DU: hu:h ne diyelim >o zaman< =customers-  
*what shall we say then*

14 (1.7)

15 wait for fshop-f ((laughs))  
+ turns gaze to SA

16 SA: buy nothing at all desek bu sefer de direk (.) çok  
*if we say "buy nothing at all", this time then directly too*  
+ DU laughs  
+ DU shifts gaze to GE and then back to her paper

17 (.) [şey ]  
*like*

18 GE: [((clicks tongue))]

19 SA: çelişcek =çok [belli olcak di mi?]  
*it will contradict, it will be too obvious won't it?*  
+ shifts gaze to GE and then back to her paper

20 GE: [olmaz evet. ]  
*it won't be possible yes*  
+ shifts gaze to SA and then quickly back to her paper

21 (1.5)

22 DU: °dur bakıyım ne denir buraya°  
*let me see what can be said here*  
+ turns over the page to look at the text

23 ((DU and SA produce thinking face with DU's hand on forehead and SA's hand over mouth and look at the text for 9.5s while GE changes gaze between the text and the item by turning over the page))

24 DU: people feel better olur mu?  
*would it be okay?*  
 + SA and DU turn over the page to look at the item  
 + GE turns over the page to look at the text and then back to the item  
 + SA turns over the page to look at the text

25 (2.0)  
 26 o da doğru [olur a↑ma (.) doğru olur. ]  
*but it would also be correct, it would be correct*  
 + turns over the page to look at the text

27 SA: [o da ama doğru işte bi şey geneline]  
 28 baktığınız zaman parçanın  
*but it is also something right you see when you look at the overall text*

29 ((DU and SA display thinking face with DU's hand on forehead and SA's hand over mouth for 21s as they all look at the text and change their gazes between the text and the item by turning over the page))

30 DU: şöyle diyelim (.) customers (.) pre↑fer (.) u:hm  
*let's say this*  
 + DU and GE look at the item on their papers  
 + SA turns over the page to look at the item

31 (0.5)  
 32 ↑natural (0.4) and  
 + all turn over the page to look at the text

33 (2.0)  
 34 natural-  
 35 (1.5)  
 36 music gibi bişiy (.) desek  
*what if we say something like that*  
 + shifts gaze to SA

37 ((SA and GE turn over the page to look at the item for 2.8s))

38 SA: ama muzak (.) soundscape olduğun↑da (.) onlar  
 + DU looks at the item on her paper

39 müzik [tercihi yapmış olmu][yo ki:↓  
*but when it is "muzak soundscape", they do not happen to have made a music preference*  
 + shifts gaze to GE and then DU twice repeatedly

40 GE: [kullanıldığında= ]  
*when it is used*

41 DU: [=kullanıldığında  
*when it is used*

42 GE: ° hnm °

43 DU: evet (.) [°ol [muyo o da° ]  
**yes it is also not possible**  
+ turns over the page to look at the text

44 GE: [°tercih [yapmıyo° ]  
**they do not make a preference**

45 SA: [yani customer]  
+ turns gaze back to her paper

46 müzik tercihi yapmıyo ki [zaten ]  
**well the customer does not make a music preference anyway**  
+ GE turns over the page to look at the text

47 DU: + turns over the page to look at the text  
[dur bakayım]  
**let me see**

48 (0.2)

49 °daha nasıl (.) zorlayalım°  
**how else shall we force it**

50 (0.2)

51 SA: °customers prefer-°  
+ GE looks at the item on the other page

52 (1.7)

53 MORE expensive products diyelim  
**let's say**  
+ shifts gaze to DU  
+ DU shifts gaze to SA

54 (1.0)

55 öyle bişey yok.  
**there is no such thing**

56 (0.2)

57 DU: [nasıl?  
**sorry?**

58 GE: [customers-  
+ DU turns gaze to the item on her paper

59 (0.2)

60 SA: pre↑fer=  
+ shifts gaze to GE  
+ GE shifts gaze to SA

61 GE: =>prefer to< bu:y-  
+ turns gaze back to the item on her paper

62 SA: huh prefer to buy more expensive ya da (.) expensive  
**or**  
+ turns gaze back to her paper and then to GE and DU twice repeatedly

63 products  
+ GE looks at the text on the previous page

64 GE: olur  
**alright**

65 DU: hu:h

66 SA: çünkü öyle bişey yok [aslında  
**because there is actually no such thing**

67 GE: [yok.  
**no**

68 (0.3)

69 SA: yani daha fazla para kazanmasını >sağlamaya çalışıyor  
**I mean it tries to enable it to earn more money**  
+ shifts gaze back to her paper with her left hand pointing at her paper  
+ DU shifts gaze to SA  
+ shifts gaze to DU

70 ama< bunu daha çok şey alsın diyerek yoksa gidip  
**but this is by telling him/her to buy more things,**  
**otherwise go**

71 (0.3)

72 DU: [°huh°

73 SA: [iki aynı bölümden pahalısını al demiyo=  
**buy the expensive one from the same two sections, it**  
**does not say**  
+ shifts gaze to GE + turns gaze to DU  
+ DU shifts gaze to GE and then back to SA

74 GE: [en pahalısını-  
**the most expensive one-**  
+ shifts gaze to SA and then to the item on her paper

75 DU: [alma  
**don't buy**  
+ GE nods head

76 SA: [=öyle bi bişey [yok  
**there is no such thing**  
+ GE nods head  
+ turns gaze back to her paper

77 DU: [demiyo [↑demiyo  
**it does not say it does not say**

78 tamam evet o zaman (.) net olur  
**okay yes then it would be clear**  
+ shifts gaze to the item on her paper  
+ SA crosses out the problematic option

79 (0.2)

80 SA: <sup>1#</sup>customers  
+ GE shifts gaze to SA and then quickly back to the item on her paper

81 (0.4)

82 pre↑fe:r (.)

83 GE: °huh-hu:h°

84 SA: to bu:y (.)

85 GE: °huh-huh°

86 (0.2)

87 SA: expensive products<sup>#1</sup>

#1 (lines 80-87) SA writes the new option as a replacement of the problematic one on her paper while reading it aloud

As all the three participants are looking at their exam papers and produce thinking face with DU and SA's hands over their mouth and GE's finger tapping on her forehead in search of an alternative for the problematic option, the tester SA offers the first suggestion for a change with an elongated hesitation marker in turn-initial position. She suggests replacing the trouble source "*fast*" in the option with its opposite "*slow*". However, turning over the page and pointing her left hand to the fourth paragraph in the reading text, she immediately challenges her own suggestion in the same turn-at-talk in lines 1, 2 and 4 by showing evidence from the reading text, that is the presence of the suggested information in paragraph four, which would also make the option correct. As a result, SA backs down and withdraws her suggestion through account-giving and thus, the realization of such evidence against her suggestion leads to self-cancellation as a form of self-initiated self-repair. During SA's turn-in-progress, DU overlaps and confirms SA in line 3 by using anticipatory completion (Lerner, 1996a) and completing SA's ongoing TCU of account-giving which she initiated in line 1. In line 5, GE also displays agreement and affiliation with SA's oppositional stance towards her suggestion in an overlapping turn at a turn-final position by clicking her tongue and providing the negative response that the suggested change cannot be possible.

After a gap of silence as all the participants are thinking of an alternative option, DU takes over in line 7 and provides the second suggestion for the problematic option, which is its negative version (i.e., *don't do shopping*). Following a lapse of 1.7 seconds with no uptake from the other participants, DU then provides a justification for her suggestion by claiming that it would eliminate the option. However, SA displays disagreement with DU in line 11 by expressing dispreference and challenging her suggestion with an orientation to a testing principle. According to her, such a change would make the option the only negative option, violating one of the adopted testing principles and thus, the tester's situated understanding of how a test should be. While challenging DU's suggestion, SA shifts her gaze to DU and then back to her paper through the end of her unit of talk. Such an account based on a testing principle is not questioned further and directly accepted by the others in the group, which shows that the participants orient to SA's account as common knowledge



regarding the testing principle for multiple-choice item construction which states that “options should be similar in structure” (Fulcher, 2010, p.172). Examining and checking the item after SA’s objection with her gaze at her paper, DU explicitly shows agreement in line 13 through an elongated acknowledgement token (hu:h) while SA and GE are already looking at their exam papers in search of another alternative.

Then DU uses a preface with an inclusive expression of suggestion (ne diyelim/*what shall we say*), which projects the upcoming suggestion and decision making process for the problematic option. Immediately following her preface, DU rushes to formulate a suggestion on spot as evident in her intra-turn cut-offs and pause as well as the articulation of her suggestion with a jokey voice followed by her laughter just as she shifts her gaze to SA. She formulates a whole new option by changing the content of the original option. However, her jokey voice and laughter indicate that the suggestion is not taken seriously and not preferred by DU herself and is uttered spontaneously while trying to formulate an appropriate option. Thus, the other participants also do not orient to her suggestion and focus on the test material without changing their gazes.

In line 16, with an inclusive language (desek/*if we say*), SA provides the fourth suggestion (i.e., buy nothing at all), which also fully changes the content of the problematic option by giving it a negative meaning. Yet in lines 16 and 19, she once again immediately challenges her own suggestion based on the account that this would contradict the information in the text, making the option too obvious to eliminate. As a result, SA backs off and withdraws her suggestion through account-giving, leading to self-cancellation. During SA’s turn-in-progress, GE overlaps in line 18 and shows disagreement with the suggestion by clicking her tongue. Upon this, while proceeding with her explanation, SA orients to GE with a shift of gaze to her and uses a tag question (di mi?/*won’t it*) to mobilize support for her account against her suggestion in turn-final position (Heritage, 2012). Directing her gaze towards SA and maintaining mutual gaze with her, GE displays agreement and affiliation with SA’s stance in line 20 in an overlapping turn with her confirmation request by pointing out the unacceptability of the suggested option (olmaz/*it won’t be possible*) and using a compliance token (evet/*yes*). With their gazes shifted back to their exam papers, the participants all continue to seek new alternatives in silence.

Then in line 22, DU applies a preface again for the upcoming suggestion-making process which is embodied as she turns over the page to look at the text. DU's preface indicates the transition back to their search for an alternative. A gap of silence for 9.5 seconds ensues as all the participants are looking at their exam papers thinking of an alternative option. During this time, DU and SA produce thinking face with DU's hand on her forehead and SA's hand over her mouth as they look at the reading text and GE changes her gaze between the text and the item by turning over the page. Then in line 24, just as SA and DU turn over the page to look at the item and GE does so to look at the text, DU takes over to offer the fifth suggestion (i.e., *people feel better*) followed by a confirmation request. At this point, SA turns over the page again to look back at the text and a gap of silence ensues with no uptake from the other participants as they are thinking of the offered suggestion. Turning to the text on the previous page, DU then displays dispreference herself in line 26 by challenging her own suggestion and claiming twice repeatedly that it would also make the option correct, which would retain the problem in the item. As a result, she withdraws her suggestion through account-giving and employs self-cancellation in the form of a self-initiated self-repair. In an overlapping turn, SA shows agreement and alignment with DU's challenging of her suggestion by employing a full modified repeat (Kim, 2002; Stivers, 2005) of DU's assertion and providing further account and evidence for the correctness of the offered option with reference to the reading text as a whole.

A long gap of silence proceeds for 21 seconds when all the participants look at the text and change their gazes between the text and the item by turning over the page. DU and SA also produce thinking face with DU's hand on her forehead and SA's hand over her mouth. In line 30, DU initiates a turn with a preface by using an inclusive expression of suggestion (*şöyle diyelim/let's say this*) which sets a ground for an upcoming suggestion. She then formulates a new alternative on spot between lines 30 and 36 (i.e., "*customers prefer natural music*") as signalled by her frequent pauses, an elongated hesitation marker, the repetition of the word "*natural*" twice during her turn-in-progress and her change of gaze between the item and the reading text on the previous page. The pitch rise in word-final position of "*pre↑fer*" and in word-initial position of "*↑natural*" is also an indication that the idea has come to her mind just then. However, she presents her suggestion just as an example to consider and leaves it to the other participants to confirm it by using an inclusive language in line 36 (*gibi bişiy (.) desek/what if we*

*say something like that*) with a shift of gaze to SA. As DU formulates the suggestion, GE and SA look at the item on their exam papers and then orient to the reading text by turning over the page simultaneously with DU. This is followed by a lapse of 2.8 seconds as DU is looking at SA, and SA and GE both turn over the page to look back at the item, thinking about DU's suggestion.

SA launches a *but*-prefaced turn in line 38, which projects her objection to DU and displays dispreference and disagreement by challenging her suggestion. In doing this, she provides account and evidence for the unacceptability of the suggested option as she directs her gaze towards GE and then to DU twice repeatedly while they are looking at their exam papers. She explains that when Muzak soundscapes are used, which is the stem, people are not in a position of making a music selection, which refers to the information suggested by DU. In this way, she shows that the suggested option does not meaningfully fit well with the stem. In an overlapping turn, GE orients to the stem which SA has referred to in her prior turn and articulates the Turkish version of the verb “*are used*” given in the stem. With no time lapse, DU repeats the same Turkish verb in line 41 by overlapping SA's turn at a turn-final position. Thus, making reference to the stem, which is the evidence presented by SA, DU and GE both think about SA's account regarding the connection between the suggestion and the stem. Then between lines 42 and 44, DU and GE support and endorse SA's oppositional stance towards DU's suggestion with a soft tone of voice as they keep their gazes on the test material. DU does it by using the compliance token “*evet/yes*” and then confirming that the suggestion is not possible while turning over the page to look at the text, whereas GE does it by using the acknowledgement token “*°hnm°*” and then employing a full modified repeat (Kim, 2002; Stivers, 2005) of SA's final words of her prior turn (*°tercih yapmıyo°/they do not make a preference*), which demonstrates her understanding of SA's account and confirms it. In lines 45 and 46, shifting her gaze back to her exam paper, SA repeats her account, highlighting the point that the customers do not make a music preference, which once again displays her dispreference with regard to DU's suggestion. During her turn-in-progress, SA and GE both turn over the page to look at the text while DU is already looking at the text. Thus, they all orient to the reading text in search of another alternative option.

In line 47, DU applies a preface with an inclusive language, which projects the upcoming suggestion-making process and signals the transition to their search for another alternative. While they are all looking at their papers, SA offers a new suggestion in lines 51 and 53 using an inclusive expression of suggestion (*diyelim/let's say*). She produces the sentence spontaneously by orienting to the first component of DU's prior suggestion (*°customers prefer°*) with a soft tone of voice and then stopping for 1.7 seconds in an attempt to find a different way of completing this sentence, reformulating DU's suggestion. Thinking of a new point, SA completes the sentence that starts with "*customers prefer*" and offers her suggestion (MORE expensive products) by uttering the first word "MORE" at a higher volume with an emphasis on it and maintaining mutual gaze with DU. After a gap of silence with no uptake, SA gives an account for her suggestion in line 55 by explaining that there is no such information in the reading text, which would make the option appropriate as a distractor.

Upon this, DU initiates a repair action in line 57 by asking SA to reproduce her suggestion. In an overlapping turn with DU, GE also initiates a repair action in line 58 by articulating the first word "customer" in the option suggested by SA and then, stopping to construct a DIU (Koshik, 2002) in order to elicit SA's repetition of her suggestion. Engaging in mutual gaze with GE, SA completes GE's DIU in line 60 by uttering the following word "pre↑fer" in her suggested sentence with a pitch rise in turn-final position. Just at the moment when SA articulates the word, GE takes the turn again, cutting off SA's utterance as she repeats SA's "*prefer*" and then reformulates this part of her suggestion in a fast pace by extending "*prefer*" with a second verb (i.e., *to bu:y*) that was not uttered by SA previously. At this point, GE constructs another DIU (Koshik, 2002) by producing this verb in an elongated form and then stopping in order to elicit the rest of the sentence from SA. Just prior to the completion of her turn, GE shifts her gaze back to the item on her exam paper. Then in line 62, SA shows agreement with GE's reformulation of her suggested sentence with "*to buy*" through an acknowledgement token and confirms it by repeating GE's prior turn (i.e., *prefer to buy*). Directing her gaze towards her paper and then to GE and DU twice repeatedly, she completes the sentence by repeating her suggestion (i.e., more expensive products) with an emphasis on "more". However, this time she presents it with two alternatives, that is either with or without "*more*". Thus, GE and SA jointly produce the new option through collaborative turn completions.

In lines 64 and 65, GE and DU respectively show preference and agreement with SA as they are looking at their papers. In doing this, GE uses an acceptance token, whereas DU produces an elongated acknowledgement token. Then SA repeats her account in line 66, reasserting that the information presented in the new option is not available in the reading text. GE confirms SA in line 67 through a partial modified repeat of SA's account (Kim, 2002; Stivers, 2005) and thus, displays an affiliative stance. Following a gap of brief silence, with her left hand pointing at her paper, SA initiates a non-minimal post-expansion sequence (Stivers, 2013) where she provides further elaboration of her prior account between lines 69 and 76 by presenting more concrete evidence with reference to what is said in the reading text about the issue of money. Considering the content of the reading text, she explains that Muzak soundscapes help earn more money by making people buy more products but not more expensive ones, which is a point that makes the new option incorrect as a distractor and solves the problem in the item. At certain points during her turn-in-progress, SA orients to the other participants by shifting her gaze to DU and maintaining mutual gaze with her and then to GE before turning back to her paper just prior to the completion of her account. In addition, DU overlaps in line 72 and exhibits further agreement with SA through an acknowledgement token that is uttered with a soft tone of voice.

Maintaining mutual gaze with SA, GE and DU also both engage in anticipatory completion (Lerner, 1996a) respectively during SA's account-giving through collaborative turn sequences (lines 74 and 75). Anticipating what is to come next, they complete SA's ongoing turn in line 73 with short phrases (*en pahalısnı/the most expensive one*, *alma/don't buy*) while she is already producing her account. In line 77, DU overlaps SA's turn again where she uses a partial modified repeat (Stivers, 2005) of SA's prior turn twice (*demiyo/it does not say*) with a rising intonation on the first syllable of the second one. GE's head nod also accompanies DU and SA's turns-at-talk between lines 75 and 77. In these ways, DU and GE display their uptake of the issue at hand, confirm SA and affiliate with her stance (Hayashi, 2013). The process of account-giving for the suggestion is followed by DU's further display of a stance of acceptance of the suggestion through her use of two compliance tokens and positive assessment with regard to the suggestion (*net olur/it would be clear*) in line 78. Upon their acceptance of the suggested change through collaborative account-giving and display of affiliative stance and agreement, the participants all turn to their papers as SA crosses out the problematic option and writes the

new option as a replacement of the problematic one on her exam paper. While writing the option down, she reads it aloud word by word between lines 80 and 87 with a stretching sound on “bu:y” and “pre↑fe:r” and a final pitch rise on “pre↑fe:r”. As SA reads aloud the option and stops after the articulation of each word accompanied by her writing, GE shows agreement through a quietly uttered acknowledgement token in lines 83 and 85.

Through a collaborative process of suggestion and decision making, extract 4b demonstrates how the participants interactionally deal with the problem identified in the first phase of the IR interaction with respect to the distractor option “b” in a multiple-choice reading comprehension item. In order to make this option incorrect as a distractor, the participants offer some suggestions until they decide on one of them. It is seen that one suggestion paves the way for another suggestion in the talk-in-interaction as the participants are constantly in search of a better alternative. In this particular extract, seven different suggestions are offered to reconstruct option “b” and the group decides on the final one (i.e., *customers prefer to buy more expensive products*). The first two suggestions are related to making a small change in the already available option (i.e., changing the adverb in the problematic option into its opposite, making the problematic option negative). However, when these are rejected based on the account that the first suggestion would still make the option correct, retaining the problem in the item and the second one would violate the testing principle that the “options should be similar in structure” (Fulcher, 2010), this initial process indicates that the option requires a greater change in terms of its content. As a result, in the upcoming five suggestions, the participants produce sentences that deviate from the original option (i.e., *customers do shopping very fast*) by completely changing its content (e.g., *customers buy nothing at all, people feel better, customers prefer natural music*). Only when the content of the offered alternative is not challenged, it is eventually accepted as in the case of the final suggestion. Thus, with the formulation and evaluation of each suggestion, the participants make an attempt to produce a better one that does not have the weaknesses of the previous ones. This shows that the articulation of every new idea that comes to mind and its discussion in and through talk is important in improving the suggestions and paving the way for better alternatives that eventually lead to change. This phase, thus, demonstrates the role of interaction in testing all the possibilities in the decision-making process until a mutually acceptable change is made and the problem is overcome.

As in extract 1b, a collaborative process is followed via some interactional actions while seeking a change for the option and the suggestions are not simply listed one by one as distinct entities. First of all, the act of challenging a suggestion through account-giving is the most apparent interactional action emerging in the phase of SC in this extract since of the seven suggestions, five of them are challenged and decided not to be acted on. The tester SA and the reviewer DU are active in doing this as they challenge either their own suggestions, leading to self-cancellation or the others' suggestions by providing account and evidence for their unacceptability as a replacement of the problematic option. Thus, while challenging a suggestion, the participants present concrete evidence and orient to the testing principles for multiple-choice item construction, which eventually lead to agreement and confirmation in the group with respect to the inappropriateness of the offered option as a distractor. For example, in response to the offered alternatives, the accounts that (1) the suggested change would make the option the only negative one (suggestion 2), (2) the information in the suggested option is present in the reading text, making the option correct (suggestions 1 and 5), (3) the suggested option does not meaningfully fit well with the stem (suggestion 6), and (4) the information in the suggested option contradicts that in the reading text, making it too obvious to eliminate (suggestion 4) are all related to the violation of certain testing principles for multiple-choice item construction and thus, are agreed by the group members without being further questioned and counter-challenged. In this respect, the act of challenging a suggestion can be considered to make an important contribution to the validation process since it prompts a critical and careful thought on the suggestion in and through talk (Fulcher & Davidson, 2007). As a result, such a collaborative way of understanding the suggestions leads to the final decision with respect to the change to be made in the problematic item "in open discussion, through a process of dialogue and disagreement" (Fulcher & Davidson, 2007, p.21).

With respect to the act of challenging a suggestion, it is employed by either the participant who offers the suggestion (i.e., suggestions 1, 4, 5) or the other participants in response to the suggestion (i.e., suggestions 2, 6) where they explain why the suggestion would cause trouble. In this extract, self-cancellation of the suggestion is, thus, one of the most frequently occurring interactional features. It occurs through account-giving against the suggestion in the form of a self-initiated self-repair. However, in the case of the third suggestion (i.e., *customers wait for shop-*), just as DU is in the process of creating the sentence, she uses a non-verbal account (i.e., jokey voice and laughter) to signal her own dispreference with her

suggestion and her self-cancellation. As a result, the other participants just disregard the suggestion, continuing with their search for an alternative. Self-cancellation of a suggestion shows that the participants formulate the suggested sentence on spot as soon as the idea comes to their mind without thinking too much on it and hence, immediately after producing the sentence, they back down and withdraw their own suggestion by challenging it. In a way, through those suggestions, they reflect their online thinking process while looking at their exam papers in search of an alternative for the problematic option. Only after producing the sentence, they come to the realization that the offered sentence was not appropriate. However, even though those suggestions are offered without too much consideration beforehand, this seems to have interactional implications as they enable the participants to approach the item with a critical eye, leading to the emergence of new ideas and discussions in the course of solving the problem in the item.

The second interactional action appearing in this extract is account-giving for the suggestion before a decision is made. The analysis of the micro-details showed that different from the way the other suggestions are interactionally framed, the final suggestion that is approved by the group (i.e., *customers prefer to buy more expensive products*) is found to follow a different sequential path in that the tester SA who produces this suggestion provides account and evidence for why it would be an appropriate option, which seems to be an important step for its acceptance. Although the participants' display of agreement with the suggestion does not project an expansion, she even initiates a non-minimal post-expansion sequence where she provides further elaboration of her account by presenting more concrete evidence from the reading text to prove the appropriateness of the formulated sentence. In this way, she shows that the text includes no information whatsoever about "*expensive products*" and the new option does not pose any risk. Thus, as in the other extracts, the suggestion that is justified based on evidence is likely to be accepted and put to use by the group.

Thirdly, as in extract 1b, it is seen that a suggestion that is offered by one participant can be reformulated and redesigned by another participant through the talk-in-interaction. The final suggestion, in fact, emerges as the reformulation of the previous one since SA picks the first part of the previously suggested sentence (i.e., *customers prefer*) and reformulates it by completing it with a different piece of information. In this respect, the previously offered suggestion that was rejected leads to a new idea in and through talk, illustrating the importance of interaction during the suggestion and decision making process. This is also



evident in the collaborative reformulation of the accepted suggestion through GE's repair action and use of a DIU as she adds a verb (i.e., *to buy*) to the option SA has offered while reproducing the sentence together with SA through collaborative turn completions. This shows that the suggestions are collaboratively shaped in and through talk.

Fourthly, when the participants do not challenge and self-cancel their own suggestion immediately after producing it, they happen to prompt a response and an evaluation of the presented suggestion from the other participants through the gaps of silence that follow their suggestions (suggestions 2, 5, 6, 7), their confirmation requests (suggestions 4, 5) as well as their presentation of a suggestion just as an example for the group to consider (suggestion 6). In this way, the participants downgrade their suggestions and do not assert epistemic priority as the first assessor or provider of the suggestion (Heritage & Raymond, 2005) but invite others' contributions to the evaluation process of the suggestions, sharing the responsibility for the decision. Since they all have equal epistemic access to the referent situation, that is the reading text, this allows for a careful thought on the offered options and a 'healthier' decision-making. On the other hand, considering the suggestion that is accepted, SA displays a stance of certainty towards her suggestion by uttering the word "MORE" at a higher volume with an emphasis on it, giving account for it in a declaratively-formatted turn (i.e., *öyle bişey yok/there is no such thing*), and repeating the same assertion several times in the talk-in-interaction. Her stance of certainty might have an important role in the group's uptake of and decision on SA's suggestion.

Another interactional resource emerging in the phase of SC in this extract is the use of a preface which projects the announcement of an upcoming suggestion. In this extract, DU does this four times throughout the interaction after a suggestion is challenged and rejected. In this way, she indicates the transition back to their search for an alternative and thus, ensures continuity in the suggestion-making sequence. In one instance in the case of sixth suggestion, she applies a preface as a projection of a new suggestion that she is in the process of formulating and draws the other participants' attention to it.

Finally, anticipatory completions (Lerner, 1996a) and full or partial modified repeats (Kim, 2002; Stivers, 2005) are the other interactional resources the participants use to display a stance of agreement and affiliation with an account given either for or against a suggestion. Through such collaborative turn sequences, the participants display and demonstrate their

shared understanding of the accounts and overtly confirm them. This shows that the participants all actively and critically think about the suggested option and challenge or support it through collaborative account-giving. For instance, although GE was relatively more passive in the suggestion-making process, she contributes to the interactional process by using an anticipatory completion and modified repeats of the given accounts and thus, displaying her evaluative stance with respect to the suggestion. In addition, by means of such interactional resources, the participants do not only show agreement but also claim their epistemic right to confirm an assertion, indexing epistemic authority and control over the matter at hand since they all have equal epistemic access to the reading text (Heritage & Raymond, 2005; Stivers, 2005). In this way, they “assert the primacy of the second speaker’s rights to make the statement” (Stivers, 2005, p.131). Besides anticipatory completion and modified repeat, presenting further evidence is found to be another way of displaying agreement with respect to the challenging of a suggestion. SA does this in the case of the fifth suggestion by making reference to the overall content of the reading text. In this way, she supports and endorses the remark already made with respect to the suggestion. Finally, offering a positive assessment of a suggestion as DU does in line 78 is an interactional resource used to display an affiliative stance with a suggestion and thus, to support a decision made with respect to the change in the item.

This extract shows an orientation to the tester roles and responsibilities in the IR interaction. To begin with, the tester SA mainly plays the role of challenging a suggestion by quickly providing account and concrete evidence from the reading text and by orienting to the way the stem is designed as well as the testing principles that display her situated understanding of how a test should be. This signals her right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item as the writer and thus, the ‘owner’ of the item. In addition, SA orients to her role as the tester who is accountable for the test by initiating a non-minimal post-expansion sequence where she provides further elaboration of her account for her suggestion although the group’s agreement with the suggestion does not project expansion. In this way, before marking the change on her exam paper, she ensures shared understanding and makes sure that the decided change indeed makes option “b” incorrect as a distractor and the problem is overcome. This is important for the tester due to accountability issues in testing because in any case of criticism, as the tester of the institution, she will have to explain to the stakeholders why the item is designed in such a way. For the same reason, SA reads aloud the new option word by word as she is writing it

down on her paper and elicits GE's acknowledgement. This can be considered as the tester's attempt for a final check and confirmation request for the change in the item before it is finalized, which once again solicits mutual understanding with the other participants.

Despite an orientation to the institutional role distribution of the participants, they happen to share the responsibility for the change process in the problematic item. This is most evident in both the reviewers and the tester's constant search and online formulation of an alternative to overcome the problem in the item. For instance, while formulating a suggestion spontaneously, the participants make their online thinking process clearer by means of elongated hesitation markers, intra-turn cut-offs, frequent pauses, repetition of the same word and change of gaze between the reading text and the item. Similarly, while the participants are searching for an alternative or thinking about an offered alternative in silence, they are seen to be all fully occupied with the test material by continuously turning over the page to look at either the reading text or the item and producing thinking face. In this respect, as they continuously look at their exam papers, the participants' gaze behaviour is another indicator of their active engagement in the task of suggesting a change for the problematic option in such a competing activity (Goodwin, 1981; Rossano, 2013). As a result, the participants all make a contribution to the interactional process by, for example, providing suggestions, challenging them, giving account, reformulating the suggestions and displaying a stance of agreement or disagreement before making a final decision. This can be explained in terms of the participants' assumption of equal epistemic state and authority since they all put themselves in K+ position as reviewers, teachers and testers who have equal epistemic access to the reading text as the main source of the test items (Heritage, 2013). Moreover, the participants' use of an inclusive language for six times throughout the interaction either while projecting a suggestion in the preface or offering a new suggestion once again ascribes an equal epistemic state to the whole group and makes reference to their shared responsibility for the change process in the item and their joint decision-making. Finally, SA's reading aloud while writing down the decided option on her exam paper signals that the change is considered as the product of the whole group and the tester is seeking final confirmation from them.

Extract 5b presents the phase of SC in the reviewing interaction on a grammar item in a multiple-choice cloze test which was problematized in extract 5a. It is to be remembered that in this extract, the two teachers MI and AV teaching at B1 plus level do the reviewing of the

item and provide their feedback to the tester ME. In the first segment of the extract (i.e., extract 5a, see Section 4.1), the reviewer MI problematized the item, which tested the students' knowledge of the past perfect tense (i.e., *had seen*), claiming that the word “*before*” used in the target sentence in the text as a clue for past perfect tense did not fit well in the sentence and was inappropriate within the given context. Thus, in the following segment of the extract (i.e., extract 5b), having already announced and agreed with the problem, the IR group finds an alternative to revise the target sentence without using “*before*” through a process of suggestion and decision making for change and improvement in the item.

### Extract 5b. Before, Segment b

- 1 MI: >mesela< [bel ]çi- (.) belki be↑fore yerine (.)  
*for example maybe instead of "before"*  
+ all look at their exam papers  
+ shifts gaze to AV and then back to her paper  
+ ME shifts gaze to MI's paper
- 2 AV: [ °huh° ]
- 3 MI: hani belki biraz da zorlamak  
*well maybe to also make a bit difficult*  
+ ME turns gaze to MI and then back to MI's paper
- 4 (0.2)
- 5 er in their ↑early-  
+ shifts gaze to AV and then to ME  
+ ME shifts gaze to MI
- 6 (0.8)
- 7 er [school  
+ ME turns gaze to her paper
- 8 ME: [educa- (.)
- 9 MI: education [time ]  
+ shifts gaze to AV
- 10 ME: [ °time° ]
- 11 MI: gibi hani öyle daha önceki eğitim hayatında  
*something like that, you know like in the previous education life*  
+ shifts gaze to ME and then back to AV
- 12 (0.5)
- 13 daha [önce- ]  
*previous-*
- 14 AV: [hani ↑bel]ki o before'a bişey ekleyelim mi  
*well maybe shall we add something to that "before"*

15 =daha (.) hani .hh  
+ MI turns gaze to her paper

16 (0.5)

17 güzel otursun [cümle içine diye mi ]  
*is it because so that you know it fits better in the sentence*

18 ME: [o zaman şey yapabili]riz.  
*then we can make it this way*

19 (0.3)

20 a few books (.) in school be↑fore ya  
*you see*  
+ reads aloud the last part of the target sentence by moving her pen along it on her paper  
+ AV follows ME by moving her pencil along the sentence on her paper

21 (0.2)

22 AV: huh-[huh ]

23 ME: [in: ] school before' ↑u (0.2) <sup>1#</sup> tamamen  
*completely*  
+ MI turns gaze to ME's paper

24 değiştirip  
*change*

25 (0.7)

26 i:n (.) earlier<sup>#1</sup> (0.2) er::  
+ shifts gaze to MI  
<sup>#1</sup> (lines 23-26) ME draws a frame around the trouble source "in school before" in the target sentence

27 MI: in their earl[y:] (.) edu[cation life]  
+ shifts gaze to AV

28 ME: [ly]  
+ nods head

29 AV: [education ]  
+ turns gaze to ME's paper and then back to hers  
+ ME shifts gaze to AV

30 MI: gibi [bişiy ] (.)  
*something like that*

31 AV: [huh-huh]

32 MI: istiyo kulağım.  
*my ear wants*  
+ turns gaze to her paper

33 AV: early years of their education gibi belki  
*something like that maybe*  
+ shifts gaze to ME  
+ MI shifts gaze to AV

34 [>öyle bişe-< huh]  
*something of the kind*

35 ME: [in the early ] years [of their education ] evet  
*yes*

- 36 AV: [huh huh-huh huh-huh]  
+ shifts gaze back to her paper  
+ MI and ME turn gazes back to  
their papers
- 37 ((ME and MI make the change on their exam papers))
- 38 ME: 2#° evet °  
**yes**
- 39 (0.3)
- 40 AV: okay.
- 41 (0.9)
- 42 burası tamam zaten hani şundan dolayı had'e  
+ ME turns gaze to AV's paper  
+ points at "before" on her paper with a pencil  
+ ME turns gaze back to her paper by  
nodding head
- 43 götürmüştü bizi #2  
**here is okay, well after all it took us to "had"**  
**because of this**  
#2 (lines 38-43) MI continues to write down the new expression on her exam paper

As all the participants are looking at their exam papers, extract 5b begins with the reviewer MI's offer of a suggestion for a change in the item. Just as MI initiates her turn with a preface that projects her upcoming suggestion for the trouble source "before" in the target sentence, the other reviewer AV uses the continuer "huh-huh" with a soft tone of voice in line 2, which supports the progress of MI's suggestion-making. In line 1, shifting her gaze to AV and then back to her exam paper, MI prefaces her suggestion by referring to the trouble source "before" in the target sentence with a rising intonation in word-final position and presents its replacement as a solution to overcome the problem in the item. Just prior to the first possible completion of MI's TCU, the tester ME shifts her gaze to MI's paper. Then MI interpolates a parenthetical insert into her as-yet-incomplete turn in line 3, where she pre-emptively provides account for the change she is to suggest next by explaining that such a change would also make the item a little difficult. This further acts as a preface projecting MI's announcement of a suggestion by pre-emptively indicating the nature of the suggestion that has not yet been delivered by MI. At this point, ME turns her gaze to MI and then back to MI's exam paper.

Following a gap of brief silence and MI's use of a hesitation marker just as she is thinking of a suggestion to replace "before" in the target sentence, she formulates her suggestion on spot between lines 5 and 9 by orienting to the other participants with a shift of gaze to them and

initially maintaining mutual gaze with ME. In doing this, she articulates the first component of her suggestion “in their ↑early” with an initial rise in the pitch of the word “↑early” and then aborts her TCU, stopping for 0.8 seconds and producing a pause-filler before uttering the following word “school” in the expression she is spontaneously formulating for “*before*”. However, just as MI articulates “school”, turning her gaze away from MI to her paper, ME overlaps and engages in an anticipatory completion (Lerner, 1996a) in line 8 by completing MI’s prior turn with “*education*” although she cuts off her utterance before fully pronouncing the word. This is met with a responsive uptake by MI in line 9 as she orients to ME’s anticipatory completion by repeating her word (i.e., *education*) and then completing it with “time” where she eventually finishes formulating her suggestion. Overlapping MI’s turn at a turn-final position just as MI articulates the final word of her suggestion (i.e., time), ME once again employs anticipatory completion (Lerner, 1996a) in line 10 by simultaneously completing MI’s turn with the same word “°time°”, which is uttered quietly. As a result, MI and ME together formulate an expression to replace “*before*” through collaborative turn completions. In addition, MI downgrades her suggestion by using the hedges “belki/*maybe*” and “gibi/*something like*” and presenting it as just one example (>mesela</for example) or possibility for the group to consider before making a decision about the change in the item.

In line 11, MI retains the floor, giving an account for her suggestion by translating the newly formulated expression into Turkish with a shift of gaze to the other participants. Just as she launches a turn again in line 13 following a gap of brief silence in order to further clarify her suggestion, AV cuts off her utterance in line 14 and takes over to offer the change of “*before*” as a solution once again to overcome the problem in the item with an inclusive expression of suggestion. She does this in a way to highlight where to make the change in the item and check her understanding of the suggested change by MI, ensuring shared understanding. Then in lines 15 and 17, AV rushes through into another unit of talk where she offers a candidate explanation for the change in the form of a confirmation request by orienting to the problem stated in the first phase of the IR interaction and explaining that the replacement of “*before*” would make it fit better in the sentence. Overlapping AV’s turn at a turn-final position, without orienting to AV’s candidate explanation, ME applies a preface in line 18 with an inclusive language, which projects the upcoming change and decision-

making process for the improvement of the item. Then in line 20, ME reads aloud the last part of the target sentence where the trouble source “*before*” is located (i.e., a few books (.) in school be↑fore) by moving her pen along the sentence on her exam paper while the other participants follow her with their gazes on the test material. AV’s following of ME’s reading aloud is embodied as she also moves her pencil along the sentence on her paper.

AV shows listenership in line 22 through the continuer “huh-huh” and ME proceeds with her turn between lines 23 and 26 where she first locates the trouble source by indicating the part of the sentence that is to be corrected and replaced (i.e., in: school before’ ↑u) with a rising intonation at its end and a stretching of the final sound of the word “in”, and then starts to articulate the new expression they have previously formulated with the stretch on the initial sound of the word “i:n”. This is embodied with ME drawing a frame around the trouble source “*in school before*” in the target sentence. During her turn-in-progress in line 23, MI turns her gaze to ME’s exam paper as ME draws their attention to the repairable in the target sentence. However, ME cannot accurately recall the new expression previously uttered by MI as she changes the adjective “early” into its comparative form “earlier” and omits the pronoun “their”. Then failing to recall the rest of the expression, she stops and uses an elongated hesitation marker while shifting her gaze to MI.

MI takes over in line 27 and performs a repair action with a shift of gaze to AV by using an embedded correction as she repeats the whole expression with the repair solutions “early:” in elongated form and “their” embedded into it. This at the same time completes ME’s prior turn, which was aborted. MI also makes a small change in the expression during her turn-in-progress as she substitutes “time” with “life”, further reformulating her own suggestion. As MI is articulating the new expression, ME displays her uptake of the repair solution by repeating the final syllable of the adjective “early” accompanied by her head nod while AV repeats MI’s “education” by turning her gaze to ME’s paper and then back to hers. In this way, the other participants exhibit agreement and affiliation with MI’s suggestion. In line 30, MI once again presents her suggestion as an example for the group to consider and downgrades it (gibi bişiy/something like that). In line 31, AV further shows agreement with MI’s suggestion through the acknowledgement token “huh-huh” in overlap with MI’s ongoing turn.



As MI turns her gaze to her paper and ME is looking at AV, AV takes the floor in line 33 and shifting her gaze to ME and maintaining mutual gaze with her, she reformulates MI's suggestion by changing the structure of the expression and substituting "life" with "years" (i.e., early years of their education). At this point, MI directs her gaze towards AV. In addition, like MI, AV also downgrades her reformulation of the suggestion through the hedges "gibi/*something like*", "belki/*maybe*" and ">öyle biş-e-*something of the kind*" and presents it as an example for the group to consider before the final decision. Overlapping AV's turn at a turn-final position, ME immediately displays her acceptance of AV's reformulation in line 35 as she inserts her expression accurately into the target sentence by adding "in the" just before it and then fully repeats the expression and uses the compliance token "evet/*yes*" in turn-final position. During ME's turn-in-progress, AV shows agreement and affiliation with ME's insertion of the expression into the target sentence by using the acknowledgement token "huh-huh" several times successively.

As a result, the participants all turn their gazes to their exam papers and ME and MI make the change on their papers by writing down the reformulated and finalized expression as a replacement of "*in school before*", which further displays their acceptance of the expression and their final decision with regard to the change in the item. Once ME finishes marking the change on her paper, she once again uses the compliance token "°evet°/*yes*" in line 38 with a soft tone of voice and AV produces the minimal expansion acceptance token "okay" in line 40, which proposes sequence closure. Finally, following a gap of silence, AV does a final check in lines 42 and 43 by providing account for the correct answer and explaining that "*before*" was the clue that led to the correct answer (i.e., *had seen*) by pointing at it on her paper with a pencil and referring to it with the recognitional demonstrative "şundan/*this*". In this way, she makes sure that the answer is still the past perfect tense after making a change in the item. During her turn-in-progress, while MI continues to write down the new expression on her paper, ME shifts her gaze to AV's paper and then back to hers by nodding her head, which acknowledges AV's account with respect to the correct answer.

Through a collaborative process of suggestion and decision making, extract 5b demonstrates how the participants interactionally deal with the problem identified in the first phase of the

IR interaction with respect to the use of “*before*” in the target sentence as a clue for the past perfect tense tested in a multiple-choice cloze test. In order to replace this word with an expression that would fit better in the sentence, the participants together formulate a rephrasing of “*before*” as prompted by the reviewer MI’s initiation of the suggestion-making process. Different from the other extracts, the participants do not offer a number of new suggestions or directly accept the first offered change as a whole but instead focus on the reformulation of one particular time expression suggested by MI as a possible replacement of “*before*” (i.e., *in their early education time*) and together work towards improving it so that it is appropriate within the given context in the text and the problem is solved. Thus, the interaction illustrates how this expression undergoes change and how it is collaboratively shaped in and through talk. Everytime the suggested expression is articulated, it becomes subject to small changes, eventually being finalized and decided on.

The spontaneous formulation of the expression and the interactional resources of anticipatory completions, collaborative turn sequences, repair, and embedded correction all play an important role in the reformulation of the suggested expression. First of all, MI’s initiation of her online formulation of the expression as evident in her aborted TCU, use of a pause-filler and the gap of silence following her as-yet-incomplete turn mobilizes ME’s anticipatory completion, which then leads to the first reformulation of the expression as MI’s “*in their early school*” becomes “*in their early education time*” through collaborative turn sequences. Then ME’s failure to recall the expression announced by MI prompts MI’s repair action and embedded correction, which in turn leads to the second reformulation of the expression as “*in their early education time*” is transformed into “*in their early education life*”. This is then finalized by AV’s restructuring and rephrasing of the expression with “*early years of their education*” and its insertion into the target sentence by ME as “*in the early years of their education*”. As a result, “*in their early school*” initially articulated just in the process of being formulated is eventually changed into “*in the early years of their education*” through a collaborative process and met with uptake from the other participants who decide on it and quickly mark it on their papers.

Another interactional action emerging in extract 5b is the participants’ account-giving for the change in the item at certain points throughout the interaction by means of different interactional resources. MI does it twice, first through her parenthetical insert where she preemptively provides account for the change she is to suggest next and second through her

translation of the expression she has just suggested into Turkish. In the case of the former, she provides a justification for the validity of such a change by explaining that it would make the item a little difficult since the use of “*before*” makes it quite straightforward. In the case of the latter, through translation, she clarifies the meaning of the newly formulated expression in terms of its function as a clue for the past perfect tense. Such an act of account-giving seems to play an important role in the acceptance of MI’s suggestion of replacing “*before*” with a time expression such as “*in their early*” since it is met with no objection but taken up by the group and collaboratively reformulated. Similarly, AV offers a candidate explanation for the change in the item in the form of a confirmation request, seeking clarification and checking her understanding of why the replacement of “*before*” is suggested as a change and a solution for the problem. By presenting her account from the perspective of the problem announced in the first phase of the IR interaction (i.e., the inappropriacy of “*before*” within the given context), she makes sure that the suggested change addresses the problem in the item with the purpose of making the time expression fit better in the target sentence and hence, ensures shared understanding in the group with respect to the change to be made. Lastly, AV’s account for the correct answer in the end acts as a final check as the change is being made in the item before moving on with the next item in the test, ensuring that the clue for the correct answer is preserved and the answer is still the past perfect tense and not any other tense.

Like in extract 2b, the suggestions provided in extract 5b are also downgraded through hedges and their presentation just as one example for the group to consider. MI and AV do this throughout the interaction when offering suggestions for “*before*”. In this way, they do not implicate epistemic priority as the first provider of the suggested time expressions (Heritage & Raymond, 2005) but invite others’ contributions to the change process and prompt their responses, sharing the responsibility for the decision. Since the participants all have equal epistemic access to the text in the test material and are in K+ position in terms of their English language proficiency as English teachers (Heritage, 2013), they might not want to be too assertive about their suggestions but rather leave room for other alternatives and reformulations, which would allow for a careful thought on the expression for “*before*” and thus, a ‘healthier’ decision-making. It is for this reason why the reformulations of the time expression initially offered by MI are welcomed in the group. On the other hand, while MI and AV’s downgrading of their suggestions paves the way for further collaborative thought on the expression for “*before*”, ME’s display of certainty while inserting the last suggested

reformulation of the expression into the target sentence in the text as well as her compliance token lead to its immediate acceptance by the group.

Finally, another interactional resource emerging in the phase of SC is the use of a preface which projects the upcoming change process. In this extract, first MI does it to initiate a new sequence after the item is problematized and to project the announcement of a suggestion that she has in mind by referring to the trouble source “*before*” as the word to be replaced and pre-emptively providing account for it. Then ME applies a preface after MI offers her suggestion in order to project the change process in the item based on MI’s suggestion and thus, to ensure continuity in the suggestion-making sequence.

This extract shows an orientation to the tester roles and responsibilities in the IR interaction. It is seen that the tester ME carefully monitors the item change process by applying a preface for the change process in the item following MI’s offer of a suggestion (line 18), locating the trouble source by reading aloud the last part of the target sentence and directing the other participants’ attention to it by drawing a frame around it on her paper, placing the suggested expressions in the target sentence accurately and modifying them where necessary to make them fit better in the sentence. In this way, she makes sure that the offered change addresses the problem identified with respect to the use of “*before*” in the target sentence and checks to see if it solves the problem satisfactorily, orienting to her role as the tester who is accountable for the test. By considering the suggested expressions in terms of their place in the target sentence, ME also directs the interaction towards the change itself and the final decision.

On the other hand, despite an orientation to the institutional role distribution of the participants, they happen to share the responsibility for the change process in the problematic item. This is most evident in all the participants’ attempts to jointly formulate an expression for “*before*” and to improve it through further reformulations in collaborative turn sequences. In addition, the participants’ use of an inclusive language (e.g., *ekleyelim mi/shall we add, yapabiliriz/we can make*) throughout the interaction and both the reviewer MI and the tester ME’s marking the change separately on their exam papers is an indicator of their shared responsibility for the change process in the item and their joint decision-making, illustrating their ownership of the test as the shared product. Furthermore, the participants’ gaze behaviour signals that they are all actively engaged in

the task of suggesting a change (Goodwin, 1981; Rossano, 2013) as they constantly look at their own or each others' exam papers throughout the talk-in-interaction. This indicates that they are all actively processing as they are suggesting and reformulating an expression as well as deciding on a change to improve the item.

Extract 6b presents the phase of SC in the reviewing interaction on an item in a note-taking listening quiz which was problematized in extract 6a. It is to be remembered that in this extract, the two teachers DU and GE teaching at B1 level do the reviewing of the item and provide their feedback to the tester SA. In the first segment of the extract (i.e., extract 6a, see Section 4.1), GE problematized one of the fill-in-the-blanks given in a lecture outline where the expected answer was “chemicals”, claiming that “physical reactions” mentioned in the listening passage could also be possible as a correct answer. Thus, in the following segment of the extract (i.e., extract 6b), having already announced and agreed with the problem, the IR group finds a way of dealing with such a possibility through a suggestion and decision making process.

#### Extract 6b. Physical Reaction, Segment b

- 1     SA: ben (.) er acaba şöyle mi? yapsak o zaman  
       ***I wonder if we shall do it this way then***  
       + DU and GE both look at their exam papers  
       + takes her copy of the exam paper, leaving the tapescript aside  
           + GE shifts gaze to SA  
                   + DU shifts gaze to SA's paper
- 2     =>size ama< uzatmış mı oluruz cümleyi  
       ***but do you think we would end up having extended the sentence***
- 3     (0.3)
- 4     [şu chemicals']ın sonu[↑na devamını]  
       ***at the end of that "chemicals", its continuation***  
       + points at “chemicals” on her paper with a pencil and then makes a dotted line next to it, pretending to be writing the rest of the sentence  
           + GE turns gaze back to her paper, nodding head
- 5     DU: [chemi↑cals     ]            [slash            ]  
       + turns gaze to her paper, pointing at “chemicals” with a pencil  
           + puts a slash next to “chemicals” on her paper  
           + shifts gaze to SA's paper and then back to hers
- 6     SA: <yaz:sam (.) [↓mı]>  
       ***should I write***  
       + shifts gaze to the tapescript

- 7 DU: [sla[sh (.) bence physical reactions  
**I think**  
+ moves her pencil along the slash on her paper  
+ shifts gaze to the tapescript and then back to her paper
- 8 GE: [huh
- 9 1# (0.5)
- 10 olabilir (.) evet.  
**possible yes**  
+ shifts gaze to SA
- 11 (0.9)
- 12 SA: anladım 2# yani onu artık  
+ turns gaze back to her exam paper
- 13 [öğrencinin (.) ya ]zacağı (.)
- 14 DU: [o da olur o da olur]  
**both are possible**
- 15 SA: >bişeye gerek kalmadan< biz #1 vermiş o[lalım ↑ki  
**I see, so without the need for the student to write something any longer, we shall be in a position to have given it ourselves so that**  
+ shifts gaze to DU  
#1 (lines 9-15) DU writes “physical reactions” next to “chemicals” on her paper
- 16 GE: [evet  
**yes**
- 17 DU: [evet  
**yes**  
+ nods head
- 18 SA: ona gitmesin ögre[nci.]  
**the student does not go for it**  
+ turns gaze back to her paper  
+ DU shifts gaze to SA’s paper and then back to hers
- 19 GE: [evet]  
**yes**
- 20 DU: yani  
**that’s right**
- 21 (0.5)
- 22 [verirse ] (.) doğru (.) o da doğru çünkü  
**if s/he gives it, it is right, because it is also right**  
+ GE shifts gaze back to her paper and then to DU  
+ SA shifts gaze to DU’s paper
- 23 SA: [şu şekilde] #2  
**like this**  
#2 (lines 12-23) SA puts a slash next to “chemicals” and then writes “physical reactions” on her paper

- 24 GE: eşanlamlısı[nı da vermiş oluyoruz =[ipucu da ve- ]  
**we would also end up having given its synonym, also clue-**
- 25 DU: + shifts gaze to SA  
[tabi °tabi ° physical [reaction ]  
**of course**  
+ shifts gaze to GE and then back to her paper
- 26 SA: [↑huh şey ↑key]'e  
+ shifts gaze to her exam paper  
and then to DU  
+ moves her  
pencil along the two alternatives on her paper
- 27 ikisini de vereyim.  
**well I shall give both of them in the key**  
+ DU shifts gaze to SA's paper
- 28 DU: tabi [tabi ]  
**of course**  
+ shifts gaze to SA
- 29 SA: [tamam]  
**okay**  
+ turns gaze to her paper
- 30 (0.5)  
31 tamam hocam  
**okay miss**  
+ GE turns gaze back to her paper
- 32 DU: yani chemicals'ı da yakalar =bak ben onu yakalamışım  
**well s/he can also catch "chemicals", look I have caught it**  
+ turns gaze back to her paper  
+ points at "chemicals" on her paper
- 33 mesela (.) er: GE öbürünü yakalamış ikisi de doğru  
**for example GE has caught the other one, both are correct**  
+ points left hand to GE  
+ GE shifts gaze to SA
- 34 (0.4)  
35 SA: tamam hocam  
**okay miss**
- 36 (0.3)  
37 DU: kimyasal reaksiyon  
+ GE turns gaze back to her paper  
+ SA shifts gaze to DU
- 38 (0.4)  
39 dese de  
**even if s/he says "chemical reaction"**  
+ GE shifts gaze to SA and then to DU
- 40 (0.4)

- 41           bunun içi[ne giriyo çünkü]  
**because it is involved in this**  
+ points at “physical reactions” written on her paper with her pencil  
+ shifts gaze to SA
- 42   SA:           [doğru oluyo       ]  
**it turns out to be correct**  
+ turns gaze to DU’s paper, nodding head
- 43           (0.2)
- 44   DU:   [o da  
**it is also**  
+ shifts gaze to GE
- 45   SA:   [aynen (.) za[ten ] [çünkü <kimyasal> şeyler  
**exactly after all because chemical things**  
+ shifts gaze to GE and then back to her paper
- 46   GE:           [tabi]  
**of course**
- 47   DU:           [fiziksel bi-  
**physical some-**  
+ shifts gaze to SA
- 48   SA:   fiziksel reaksiyon oluyo zaten  
**already happen to be physical reaction**  
+ DU nods head  
+ GE turns gaze back to her paper  
+ shifts gaze to DU and then back to her paper  
+ DU turns gaze back to her paper
- 49   GE:   [evet  
**yes**
- 50   SA:   [tamam hocam key’e ikisini de veriyim  
**okay miss I shall give both of them in the key**
- 51           (0.3)
- 52   DU:   °huh-huh°  
+ SA circles the two alternative answers written on her paper

As DU and GE are both looking at their exam papers and SA is taking her copy of the exam paper, leaving the tapescript aside, extract 6b starts with the tester SA’s preface, which projects the announcement of an upcoming suggestion for change in the item. While applying her preface, SA uses an inclusive language and the other participants orient to her as GE shifts her gaze to SA and DU directs her gaze towards SA’s paper. Before announcing her suggestion, SA rushes to interpolate a parenthetical insert into her as-yet-incomplete turn in line 2 where she pre-emptively presents a challenge that might possibly arise with respect to the change to be suggested by prompting the other participants to think about whether such a change would extend the sentence in the item in the lecture outline. Once again she uses an inclusive language in her parenthetical insert (uzatmış mı oluruz



cümleyi/*would we end up having extended the sentence*). Then in lines 4 and 6, SA articulates her suggestion by offering to write the other possible correct answer (i.e., *physical reactions*) as a continuation of the sentence where the blank for “*chemicals*” is given in the lecture outline and thus, to include it within the item. This is at the same time embodied as SA points at the answer “*chemicals*” on her paper with a pencil and then makes a dotted line next to it, pretending to be writing the rest of the sentence. During her turn-in-progress, the reviewer GE turns her gaze back to her paper accompanied by her head nod and then uses an acknowledgement token in line 8 just prior to the end of SA’s turn-at-talk, which signal her agreement with the suggestion.

Simultaneously with SA’s turn, DU overlaps in line 5 and turning her gaze to her paper and pointing at “*chemicals*” on her paper with a pencil, she provides her own suggestion, which is putting a slash next to “*chemicals*” and then writing the other possible correct answer “*physical reactions*”. Her articulation of the word “slash” is embodied with her putting a slash next to “*chemicals*” on her paper. Just prior to the first possible completion point of her TCU in line 5, DU shifts her gaze to SA’s paper and then back to hers as SA is at the same time explaining her suggestion. Upon DU’s overlapping suggestion, SA utters the final unit of her talk in line 6 in a slow pace with a falling intonation and an elongation while shifting her gaze to the tapescript. In doing this, she also downgrades her suggestion by directing it to the other participants as a confirmation request (i.e., *yaz:sam (.) ↓m1/should I write*). Just prior to the completion of SA’s turn, as DU proceeds with her turn of suggestion in line 7, she repeats her last word (i.e., *sl*ash) embodied with her moving her pencil along the slash already marked on her paper and then articulates the other possible correct answer (i.e., *physical reactions*) that is to be written next to the slash. During her turn-in-progress, DU also shifts her gaze to the tapescript on the table and then back to her paper, starting to write “*physical reactions*” next to “*chemicals*” on her paper. Following a gap of brief silence, GE displays agreement with DU’s suggestion through her use of an acceptance and a compliance token with a shift of gaze to SA.

Shifting her gaze back to her exam paper, SA takes a turn in line 11 with a claim of understanding with respect to DU’s suggestion. She then clarifies her suggestion and provides account for it by offering to include “*physical reactions*” in the item without the need for the students to write it down in the blank as an answer with a shift of gaze to DU

and explaining that such a change will prevent the students from going for “*physical reactions*”. In doing this, she uses an inclusive expression of suggestion (*biz vermiş olalım/we shall be in a position to have given it*). Her turn-at-talk is also embodied as she at the same time puts a slash next to “*chemicals*” as suggested by DU and then starts to write “*physical reactions*” on her paper, which displays her uptake of DU’s suggestion. While DU and SA are both marking the change on their papers, in an overlapping turn with SA’s ongoing clarification, DU also gives an account for the change in line 14 by reasserting that both alternatives can be possible as correct answers. Through the end of SA’s account, DU and GE both show agreement in lines 16 and 17 through their simultaneous use of the compliance token “*evet/yes*” accompanied by GE’s head nod. GE repeats her compliance token in line 19 just prior to the completion of SA’s turn. Then as SA directs her gaze back to her paper, continuing to mark the change on her paper and DU shifts her gaze away from SA’s paper to her own, DU takes over in line 20 where she reformulates her prior account for her suggestion with respect to the correctness of both answers. At this point, GE shifts her gaze back to her paper and then to DU and completing her noting of the change on her exam paper, SA directs her gaze towards DU’s exam paper.

GE takes over in line 24 to provide a further justification for the change with a shift of gaze to SA. She points out the benefits such a change would bring by explaining with an inclusive language that this would also allow them to give the synonym of the correct answer, which would, in turn, function as a clue for the students. During her turn-in-progress, shifting her gaze to GE and then back to her paper, DU immediately acknowledges GE with a marked confirmation (*tabi/of course*) (Stivers, 2011) and repeats the other possible correct answer “*physical reaction*” which GE has oriented to. However, GE cannot complete her turn when SA cuts off her utterance in line 26 to perform a repair action with respect to her understanding of the change suggested by DU. The round of account-giving for the change leads to SA’s realization that what DU suggested was in fact different from what she suggested and intended to do in the item. While DU has suggested giving the two alternative responses in the answer key as acceptable and correct answers, SA has offered to include the information about “*physical reactions*” within the item so that the students do not have the chance of writing it in the blank as a possible correct answer. As a result of this misunderstanding, SA launches a repair action with a turn-initial change-of-state token (Heritage, 1984b) uttered with a pitch rise and offers a candidate understanding (Antaki,

2012; Heritage, 1984b) of DU's suggestion by shifting her gaze to DU and soliciting confirmation that it is the key where the two possible correct answers are to be given with an emphasis and an initial pitch rise on the repair solution " $\uparrow$  key". This is embodied as SA moves her pencil along the two alternative answers written on her paper. This is received with DU's marked confirmation in line 28 while shifting her gaze to SA, which displays her certainty with respect to her suggestion. Turning her gaze to her paper, SA uses the minimal expansion acceptance token "*tamam/okay*" twice followed by the address term "*hocam/miss*" in turn-final position.

Following SA's repair action, DU initiates a non-minimal post-expansion sequence in line 32 for further elaboration of her suggestion (Stivers, 2013) by orienting to the announced problem that both answers could be possible as correct and ensuring that the suggested change addresses this problem. Turning her gaze back to her paper and pointing at "*chemicals*" on her paper, she once again reformulates her prior account with respect to the correctness of both answers. In doing this, she points out that the students can catch the expected answer "*chemicals*" as she has done but they can also go for "*physical reactions*" as the other reviewer GE has done by pointing her left hand to GE. At this point, GE shifts her gaze to SA, who then displays a stance of acceptance of the change suggested by DU in line 35 by addressing DU with "*hocam/miss*" again in turn-final position. Following a gap of brief silence, DU proceeds with her turn of elaboration between lines 37 and 47 where she provides further account that "*chemical reactions*" is a part of "*physical reactions*". This is embodied with DU pointing at "*physical reactions*" written on her paper with a pencil. In this way, she gives a justification for the possibility of the other alternative within the context in the item. During her turn-in-progress, GE and SA both shift their gazes to DU and SA overlaps in line 42 to employ an anticipatory completion (Lerner, 1996a) by turning her gaze to DU's paper accompanied by her head nod and completing DU's ongoing account with the remark that both alternatives can be correct. As a result, DU directs her gaze to SA while at the same time continuing with her turn and then completes the rest of her account by further relating "*chemicals*" to "*physical reactions*" with a shift of gaze first to SA and then to GE.

Simultaneously with DU's turn, SA displays agreement and affiliation with DU's stance through the use of the compliance token "*aynen/exactly*" in line 45 and then in

alignment with her turn, she produces an account that is complementary to DU's explanation by further commenting on the close connection between "*chemicals*" and "*physical reactions*" with a shift of gaze to GE and then back to her paper. Meanwhile, GE overlaps in line 46 and acknowledges the other participants' stance through a marked confirmation (*tabi/of course*), showing agreement with their account for the suggested change in the item. Upon SA's overlapping turn, DU aborts her final TCU in line 47, cutting off her utterance and SA proceeds with her explanation in line 48 by shifting her gaze to DU, at which point DU exhibits agreement with a head nod. As they all turn their gazes back to their papers, they come to a final decision about the change in the item and propose sequence closure. In doing this, GE conveys an affiliative stance via the compliance token "*evet/yes*" in line 49 and SA shows acceptance of the change through her acceptance token "*tamam*" followed by her reannouncement of the decided change, that is including the two possible answers in the key. This is received with DU's acknowledgement in line 52 as SA circles the two alternative answers written on her paper. As a result, the group decides on DU's suggestion of giving the other possible correct answer in the key rather than embedding it within the sentence in the item.

Through a collaborative process of suggestion and decision making, extract 6b demonstrates how the participants interactionally deal with the problem identified in the first phase of the IR interaction with respect to the possibility of "*physical reactions*" as a correct answer besides the expected answer "*chemicals*" in one of the fill-in-the-blanks given in a lecture outline for a note-taking listening quiz. In this extract, two different suggestions are offered simultaneously to overcome the problem, which are (1) including the other possible answer "*physical reactions*" in the item as a continuation of the present sentence, and (2) giving both alternative answers in the answer key separated by a slash. They both aim to make "*physical reactions*" appear together with "*chemicals*" so that it is not overlooked as a correct answer. What makes these two suggestions different is the place of the change. While the former refers to a change in the item itself, aiming to prevent the students from going for "*physical reactions*", the latter is about the organization of the answer key, allowing the teachers not to disregard "*physical reactions*" as a correct answer and hence, avoiding any objections from the stakeholders. However, the simultaneous announcement of these two suggestions gives rise to a misunderstanding since the participants at first happen to orient to these suggestions as the same change in the item. Thus, SA regards DU's suggestion as an offer for the format of the change she is suggesting; thereby, including the

two alternative answers in the item but separating them by a slash as suggested by DU. The main interactional action that plays a role in clearing up this misunderstanding is all the participants' account-giving for the change from different perspectives. Such a collaborative process of account-giving and discussion leads to SA's repair action and realization that they have suggested two different changes to solve the problem and ensures shared understanding in the group with respect to the change to be made in the item. As a result, the group orients to DU's suggestion and eventually, decides to give the two possible answers in the key of the test.

Regarding the interactional act of account-giving for the suggestions, it seems to be a predominant feature of the interaction in extract 6b. As it is already mentioned, it plays an important role not only in clearing up a misunderstanding with respect to the suggested changes and ensuring shared understanding in the group but also in evaluating and discussing the change critically from different perspectives and arguing for its validity. Throughout the interaction in extract 6b, the participants provide account for the change in terms of (1) its role in overcoming the announced problem, (2) the students' approach to the item, (3) its additional benefit for the students, (4) the close connection between the two alternative answers, and (5) the reviewers' test-taking experience while listening to the lecture and taking notes. Even when a mutual understanding is reached about the suggested change through repair action, which does not project an expansion, it is seen that the process of account-giving and elaboration of the suggested change continues in a non-minimal post-expansion sequence (Stivers, 2013), where the participants present further evidence for the validity of the offered change with an orientation to the announced problem and then come to a final decision. Furthermore, the interactional act of account-giving is employed collaboratively through the participants' supportive and complementary explanations, anticipatory completion and hence, collaborative turn sequences. Such a process of account-giving enables the participants to discuss the suggested changes thoroughly, which affects the decision-making process to a considerable extent.

Like in extracts 2b and 5b, the suggestion offered by SA is downgraded in this extract since it is directed to the other participants as a confirmation request and uttered with hesitation as evident in the intra-turn pause, slow pace, elongation and falling intonation marked in the final unit of her talk. In this way, SA does not implicate epistemic priority as the first provider of the suggestion while the reviewer DU is also offering another suggestion

(Heritage & Raymond, 2005) but invites others' contributions to the change process and prompts their responses, sharing the responsibility for the decision. Since the participants all have equal epistemic access to the listening passage, the notes taken on the note-taking sheet as well as the tapescript, SA might not want to be too assertive about her suggestion but rather leave room for other alternatives, which would allow for a careful thought on a solution for the announced problem and thus, a 'healthier' decision-making. On the other hand, DU's display of certainty with her suggestion and claims as evident in her marked confirmations throughout the interaction through the use of "*tabi/of course*" and her reformulations of her account lead to the immediate acceptance of her suggestion by the group. This might be the reason why SA directly orients to DU's suggestion upon her repair action and does not insist on the change she has initially offered.

Regarding the institutional role distribution of the participants in the extract, there are indications of tester and reviewer roles and responsibilities in this IR interaction. To begin with the tester roles, it is seen that the tester SA carefully monitors the item change process and takes firm steps towards the change by, for instance, pre-emptively presenting a challenge that might possibly arise with respect to her suggestion through a parenthetical insert, performing a repair action and offering a candidate understanding (Antaki, 2012; Heritage, 1984b) of DU's suggestion, and by reannouncing the decided change at the end of the talk-in-interaction as an attempt for final confirmation check for her understanding of the suggested change as correct. These orient to SA's role as the tester who is accountable for the test and her responsibility to be alert to any problems that might emerge in the IR interaction as well as in the change made in the item and to ensure mutual understanding with respect to it. As for the reviewer roles, at one point in the interaction, a reference is made to the reviewers' test-taking process by using the present perfect tense (*yakalamışım/ I have caught, yakalamış/s/he has caught*). This signals their responsibility and obligation to have taken the exam as a student would before giving their feedback on the item (Heritage, 2013).

Despite an orientation to the institutional role distribution of the participants, they happen to share the responsibility for the change process in the problematic item. This is most evident in both the reviewer DU and the tester SA's attempts to simultaneously offer a suggestion for the problem in the item and in all the participants' careful thought on the change by providing account from different perspectives. This can be explained in terms of the

participants' assumption of equal epistemic state and authority since they all put themselves in K+ position as reviewers, teachers and testers who have equal epistemic access to the listening passage as the main source of the test (Heritage, 2013). In addition, the participants' use of an inclusive language (e.g., *yapsak/we shall do*, *uzatmış mı oluruz/would we end up having extended*, *biz vermiş olalım/we shall be in a position to have given*, *vermiş oluyoruz/we end up having given*) throughout the interaction and both the reviewer DU and the tester SA's marking the change separately on their exam papers is an indicator of their shared responsibility for the change process in the item and their joint decision-making, illustrating their ownership of the test as the shared product. Furthermore, the participants' gaze behaviour signals that they are all actively engaged in the task of suggesting a change (Goodwin, 1981; Rossano, 2013) as they constantly look at their own or each others' exam papers and the tapescript throughout the talk-in-interaction. This indicates that they are all actively processing as they are suggesting and deciding on a change to improve the item.

Extract 7b presents the phase of SC in the reviewing interaction on a grammar item in a binary-choice test which was problematized in extract 7a. It is to be remembered that in this extract, the teacher CI teaching at B1 level does the reviewing of the item and provides her feedback to the tester AY. In the first segment of the extract (i.e., extract 7a, see Section 4.1), the reviewer problematized the item, which tested the students' knowledge of the modal verbs they had studied in class, claiming that the two modal verbs (i.e., *might* and *should*) given in the binary-choice could both be possible as correct answers considering the context provided in a text. Thus, in the following segment of the extract (i.e., extract 7b), having already announced and agreed with the problem, the IR group finds an alternative for the other possible correct answer "*might*" besides the expected answer "*should*" through a suggestion and decision making process.

#### Extract 7b. Should Might, Segment b

- 1 CI: oraya başka bişey ne olabilir  
*what else can be put there*  
 + produces thinking face and puts hand under her chin  
 + AY gets the textbook and searches in it
- 2 AY: <sup>1#</sup>bi dakika<sup>o</sup>  
*one minute*
- 3 (2.6)
- 4 CI: <sup>o</sup>have to<sup>o</sup> ((mouths "be"))

5 (1.0)  
6 tabi have to mu[st falan hep riskli foluyo[ zaten de  
**of course "have to" "must" and so on happen to be**  
**always risky anyway**

7 AY: [evet (.) er o o o da olmuyo aynen  
**yes it is also not possible, exactly**

8 (1.7)  
9 ilk ba[ta must gibi d[üşünmü[ştüm =ama must should'a  
10 yakın zaten =o  
11 (0.3)  
12 zorluyo. #1 (.) ba[şa da ö[ğrendikleri bi[şey yok.  
**at first I thought like "must" but "must" is already**  
**close to "should", it makes it difficult, there is**  
**nothing else that they have learned**  
+ closes the textbook by shaking her head and puts it on the table  
#1 (lines 2-12) AY searches in the textbook for alternative modal verbs studied in class while CI  
keeps her gaze on the item on her paper

13 (0.3)  
14 CI: should shouldn't çok mu basit o[đlur  
**would "should" and "shouldn't" be too easy**  
+ AY looks at the item on her exam paper

15 (1.9)  
16 AY: uhm  
17 (1.0)  
18 shouldn't ama careful o- er  
**but**  
+ points at "careful" on her paper with a pen

19 (1.3)  
20 dikkatli olmamalısın pek olmuyo [sanki. ]  
**"you shouldn't be careful" seems to be unlikely**

21 CI: [°çok sa[ç]ma di mi°  
**it is too nonsense, isn't it**

22 ((CI and AY both look at the item on their papers for  
20.0s, producing thinking face and reading some parts  
of the sentence silently))

23 2#° burda daha ne diyebiliriz ki°  
**what else can we say here**

24 (5.4)  
25 °hani-°  
**well**

26 (0.9)  
27 don't have to ↑be falan-  
28 (2.3)  
29 daha zorlar mı acaba?





56 CI: ötürü ingilizler  
57 (0.6)  
58 er:  
59 (0.4)  
60 <dikkatli olmak> ve onları korumak zorunda<sup>#3</sup> değiller  
*because of this reason the british don't have to be*  
*careful and protect them*  
+ moves her pencil along "protect" with a  
repeated alternating movement  
#3 (lines 52-60) CI moves her pencil along the target sentence while translating it into Turkish  
with "don't have to" inserted into it  
61 (0.9)  
62 ama ko[rumak zorunda<sup>#2</sup>] değiller (.) yanlış oluyo  
*but "they don't have to protect" is wrong*  
+ underlines "protect" on her paper  
+ shifts gaze to AY  
+ AY shifts gaze to CI  
#2 (lines 23-62) AY keeps her gaze on the item accompanied by a thinking face  
63 AY: [hu:h ]  
64 evet  
yes  
+ turns gaze back to her paper  
65 (1.4)  
66 CI: korumak zorunda değiller niye de[sin (.) hani (.)  
*why would one say "they don't have to protect", well*  
+ AY mouths CI's words "korumak zorunda değiller"  
+ turns gaze back to her  
paper  
67 AY: [huh-huh huh-huh  
+ nods head  
68 CI: sırf be careful olsa ↑bel[ki hani azıcık evet]  
*if it was only "be careful", maybe you know it would*  
*be a little yes*  
+ underlines "be careful" on her paper and then shifts gaze to AY  
69 AY: [korumaları gerekiyo]  
+ nods head  
70 aslında  
*actually they have to protect*  
71 CI: ama koru(.)maları gerek =>korumak zorunda< değiller  
*but they have to protect, they don't have to protect*  
+ shifts gaze to AY's paper  
+ AY crosses out "might" on her paper, then stops and produces thinking face by putting  
hand on her forehead and mouthing CI's words "korumak zorunda değiller"  
+ turns gaze back to  
her paper  
72 (0.6)

73 >e↑konomiye< yardım ediyo (.) ko-  
**it helps the economy**

74 (0.5)

75 o yüzden korumak [zorunda değiller ]  
**that's why they don't have to protect**

76 AY: [dikkatli olmaları ve]  
77 korumalarına gerek yok  
**there is no need for them to be careful and to protect**

78 (0.2)

79 CI: saçma olu↑yo  
**becomes nonsense**  
+ shifts gaze to AY

80 AY: huh-huh  
+ nods head

81 ((CI turns gaze back to her paper, AY produces thinking face and mouths a few words while looking at the item on her paper for 3.8s))

82 CI: hani bence olabilir =hem de biraz  
+ AY nods head

83 dü[şündürüyo demek ↓ki]  
**well I think it is okay, besides it means that it makes one think a little bit**  
+ shifts gaze to AY

84 AY: [bundan daha <sup>4#</sup>evet ]  
**more than this, yes**

85 CI: de mi?  
**doesn't it?**  
+ turns gaze back to her paper

86 AY: don't  
87 (1.0)  
88 °have to °#4  
#4 (lines 84-88) AY writes "don't have to" on her paper as a replacement of "might" and reads it aloud at the same time

89 (1.0)  
((lines 90 to 99 are omitted))

100 CI: hani bana şu protect'ten dolayı olmaz  
+ circles "protect" on her paper

101 gibi ge[liyo ki]  
**well to me it seems unlikely because of this "protect" and plus**

102 AY: [huh-huh]

103 CI: protect'in (.) de anlamını <bilmeleri> gerekir  
**they should also know the meaning of "protect"**  
+ points at "protect" on her paper with a pencil

104 AY: er evet  
           **yes**  
 105       (0.5)  
 106 CI: evet  
           **yes**  
 107       (0.6)  
 108       tamam o zaman  
           **okay then**  
 109       (0.4)  
 110       yani  
 111       (0.7)  
 112       bence bi problem yok  
           **so I think there is no problem**  
 113 AY: tamam  
           **okay**

Extract 7b starts with the tester CI's preface projecting the upcoming suggestion and decision making process for the problematic option in the binary-choice item. Through the end of CI's preface, AY begins to search in the textbook for alternative modal verbs studied in class while the reviewer CI is looking at the item on her exam paper with a thinking face. As their search for an alternative modal verb continues, CI articulates one of the modal verbs studied (i.e., °have to°) with a soft tone of voice in line 4. She shows her online thinking process about the validity of this modal verb as a distractor for the item by inserting it into the sentence and mouthing the following verb "be" given in the sentence. After carefully thinking about it during a gap of silence, she challenges modal verbs such as "have to" and "must" in line 6 by commenting on their inappropriacy as distractors within the given context in the item with a jokey voice since they can also be possible as correct answers in the same way as the expected answer "should". In doing this, CI presents her claim as assumed knowledge through her use of the adverb "zaten/anyway" and the marker of certainty "tabi/of course" (Stivers, 2011). In this way, before suggesting an alternative modal verb for the options, CI considers all the possible modal verbs studied in class with a critical eye on them during the suggestion-making process and ensures shared understanding regarding why they were not preferred as distractors for the binary-choice item.

In an overlapping turn with CI, AY affiliates with CI's stance in line 7 by using the compliance tokens "evet/yes" and "aynen/exactly" and acknowledging the

unlikelihood of such modal verbs through a (re)formulation of CI's turn (o da olmuyo/*it is also not possible*). Following a gap of silence, AY makes reference to her item writing process in line 9 by pointing out that "*must*" was an alternative that she thought of in the first place and then rushing to comment on the close association between "*must*" and "*should*" as the reason for her dispreference with "*must*", which further supports the stance conveyed by CI. Proceeding with her turn in line 12, AY refers to an absence of past instructional events (Can Daşkın, 2017) to indicate that the students have learned no other modal verbs. This is embodied with her closing the textbook accompanied by her head shake and putting it on the table. In this way, AY shows that she has considered all the other modal verbs studied in class (i.e., *have to* and *must*) as alternatives for "*should*", indirectly stating why she had to construct the item in such a way.

In line 14, as AY orients to the test material after leaving the textbook aside and CI sustains her gaze on her copy of the exam paper, CI offers the first suggestion, that is the negative form of the answer "*should*" with a falling intonation at the end of her turn. In doing this, CI downgrades her suggestion by presenting it in the form of a confirmation request for its level of easiness, which is tilted toward a negative reply, showing that "*shouldn't*" is already assumed to be easy by CI. By soliciting AY's confirmation about whether "*shouldn't*" would make the item too easy, CI invites collaborative inquiry into her suggestion in the course of finding a distractor modal verb. Following a gap of silence as they both think about the offered suggestion, AY takes over and challenges CI's suggestion in lines 18 and 20. Pointing at the adjective "*careful*" in the target sentence on her paper with a pen, she hesitantly claims that "*shouldn't*" is not meaningful and likely within the given context in the item. Meanwhile, she downgrades her claim through the hedge "*sanki/seems like*". Overlapping AY's turn at a turn-final position, CI displays agreement and affiliation with AY's stance by quietly offering a (re)formulation of AY's remark (Heritage & Watson, 1979), which acknowledges that "*shouldn't*" makes no sense in the sentence, and then using a tag question to mobilize support for her (re)formulation (Heritage, 2012). As a result, CI backs down and withdraws her suggestion.

Then a long gap of silence for 20.0 seconds ensues as the participants both look at the item on their exam papers accompanied by a thinking face and read some parts of the sentence silently in search of an alternative modal verb for the binary-choice item. In line 23, CI

applies a preface with an inclusive language, which projects the upcoming suggestion-making process and signals the transition to their search for another alternative. After a substantial lapse, CI makes the second suggestion (i.e., *don't have to*) in line 27, which is again a negative modal verb but this time the negative version of "*have to*". Once again, she downgrades her suggestion by presenting it in the form of a confirmation request with respect to its level of difficulty and prompts a critical and careful thought on the suggestion in and through talk. This allows a silence to elapse as AY starts to think about the new suggestion. At one point during this time, AY inserts "*don't have to*" into the target sentence and quietly reads it along with the verb expression "*be careful*" that just proceeds the blank in the sentence.

When a long gap of silence ensues with no responsive uptake from AY, CI then takes over in line 33 to provide account and evidence for her suggestion from the surrounding context in the text. Moving her pencil along the target sentence, she translates it into Turkish and explains the benefits that the crabs mentioned in the text have for the British economy. This information presents contradictory evidence for the possibility of the negative "*don't have to*" as a meaningful modal verb with the positive actions of the British, which are being careful and protecting the crabs in return of the benefits they receive from the crabs. As a result, "*don't have to*" seems to be an appropriate distractor for "*should*". After a gap of silence, CI further makes a comment on her suggestion in line 37 by referring to the students with the deictic "*onları/them*" and stating that "*don't have to*" could make the students think a little bit more on the item without directly selecting "*should*". Through the end of her unit of talk, she shifts her gaze to AY, who then shows agreement via the use of the acknowledgement token "*huh-huh*" in line 39 although she still continues to think about the new suggestion with her gaze sustained on the item and her production of a thinking face.

After turning back to her paper, just as CI reformulates her prior account with reference to the text by stating that the crabs have no harm for the British and making a circling movement on the relevant part of the text with her pencil, AY overlaps in line 44 and producing a confused face, she hesitantly challenges CI's suggestion through a confirmation-seeking polar question with respect to the possibility of "*don't have to*" within the given context. Then AY continues to think about the form-meaning relation between "*don't have to*" and the target sentence as evident in her use of an elongated hesitation

marker in line 49 and her articulation of a reason-result conjunction (i.e., *bu yüzden/because of this*) in line 51 accompanied by a thinking face.

With no time lapse, CI quickly launches a non-minimal post-expansion sequence between lines 52 and 83 where she proceeds with her account-giving to counter-challenge AY's claim against her suggestion until agreement is reached on the recommended change. In doing this, she first translates the whole sentence into Turkish with "*don't have to*" inserted into it by moving her pencil along the sentence and putting an emphasis on "*yardım/help*" and "*değiller/they don't*" in word-initial position in order to clarify and highlight the incongruity of the meaning that "*don't have to*" adds to the sentence. Just as she articulates "*korumak/protect*", she moves her pencil along the verb "*protect*" in the sentence with a repeated alternating movement to emphasize the role of "*protect*" as a clue for the unlikelihood of "*don't have to*" in the sentence. During her turn-in-progress, AY uses the continuer "*huh-huh*" in line 55, which supports the progress of CI's account-giving. CI then moves on to assert that "*don't have to*" becomes wrong with "*protect*" which is embodied as she underlines "*protect*" on her paper. Meanwhile, AY overlaps in line 63 and displays understanding of the matter at hand by using an elongated acknowledgement token and then maintains mutual gaze with CI. In line 64, turning her gaze back to her paper, AY uses the compliance token "*evet/yes*", which further acknowledges CI's account.

Following a gap of silence, with her gaze at AY, CI reformulates her remark that it is not meaningful to say "*don't have to protect*" in line 66 while AY mouths CI's prior words, inserting "*don't have to*" into the sentence and uttering it together with the verb "*protect*" in Turkish as she is still thinking about the likelihood of the suggested modal verb in the target sentence. Then in overlap with CI's ongoing turn, AY displays agreement and understanding by using the acknowledgement token "*huh-huh*" twice repeatedly in line 67 accompanied by her head nod. Turning her gaze back to her paper, CI provides further clarification in line 68 by referring to the other verb expression in the sentence just proceeding the blank (i.e., *be careful*) and indicating that "*be careful*" alone would make "*don't have to*" a little more acceptable in the sentence, which illustrates the existence of the two verbs "*be careful*" and "*protect*" in the sentence as a strong clue for eliminating "*don't have to*" in the options. This is embodied as CI underlines "*be careful*" on her paper.

Through the end of CI's turn-at-talk, just as CI shifts her gaze to AY, AY overlaps in line 69 and provides a (re)formulation (Heritage & Watson, 1979) of CI's prior account accompanied by her head nod by pointing out that the British, in fact, have to protect the crabs, which is contrary to what "*don't have to*" does in the sentence and thus, which makes it acceptable as a distractor. In this way, AY displays her understanding of CI's point with regard to her suggestion. In alignment with AY's turn, CI confirms AY through a modified repeat (Kim, 2002; Stivers, 2005) of her utterance in line 71 by shifting her gaze to AY's exam paper and then repeats the option "*don't have to*" with "*protect*" by putting an emphasis on "*değiller/they don't*" in word-initial position, comparing the meaning of the sentence with and without "*don't have to*". AY crosses out "*might*" on her paper simultaneously with CI's turn and then stops and produces thinking face by putting her hand on her forehead and mouthing CI's utterance "*korumak zorunda değiller/they don't have to protect*", which shows that AY is still thinking about the appropriacy of "*don't have to*" as a distractor.

Turning her gaze back to her paper, CI then proceeds with her turn of account-giving between lines 73 and 79 by once again repeating the meaninglessness of the reason-result relation in the sentence when "*don't have to*" is inserted into it with reference to the information in the text (i.e., >e↑konomiye< yardım ediyö/*it helps the economy*). During CI's turn-at-talk as she is translating some parts of the target sentence into Turkish to show the meaning relation with "*don't have to*", AY employs anticipatory completion (Lerner, 1996a) in lines 76 and 77. Anticipating what is to come next, she completes CI's ongoing turn by translating the verb expressions in the clause (i.e., *be careful, protect them*) with "*don't have to*" into Turkish. Just prior to the completion of CI's turn, she shifts her gaze to AY, who then displays agreement through an acknowledgement token in line 80 accompanied by her head nod. Then CI turns her gaze back to her paper and AY produces thinking face and mouths a few words while looking at the item on her paper, which still exhibits her uncertainty about the suggested modal verb despite the concrete evidence and account provided by CI.

CI takes over again in line 82 to convey her positive stance towards her suggestion "*don't have to*" and to make a positive evaluation of this modal verb by referring to the interactional process they have gone through while critically discussing it and using it as a



proof to indicate that it makes one think a little bit about the item. During her turn-in-progress, AY nods her head, displaying a stance of agreement with the suggestion and CI shifts her gaze to AY. Overlapping CI's turn at a turn-final position, AY acknowledges CI's positive evaluation of the suggested modal verb in line 84 and then starts to write "*don't have to*" on her paper as a replacement of "*might*" by reading it aloud quietly at the same time in lines 86 and 88, which shows her uptake of the suggestion and their decision on the change. Upon AY's acceptance of the suggestion, CI expresses her affiliative stance through a tag question that mobilizes support for her positive evaluation (i.e., *de mi?/doesn't it?*) in line 85 (Heritage, 2012) as she turns her gaze back to her paper.

Following the suggestion and decision making process in the phase of SC, CI initiates another non-minimal post-expansion sequence (Stivers, 2013) in line 100 to reformulate her account and to make her final comment on the change. In doing this, she once again points out that the verb "*protect*" used in the sentence makes "*don't have to*" unacceptable while circling "*protect*" on her paper. This is received with AY's display of agreement through the use of an acknowledgement token. CI then considers "*protect*" from the students' perspectives in line 103 by explaining that the students should know the meaning of this verb accompanied by her pointing at "*protect*" on her paper with a pencil. This is also met with AY's agreement in line 104 as she uses the compliance token "*evet/yes*". Finally, the sequence closes with CI's articulation of the compliance token "*evet/yes*" and the minimal expansion acceptance token "*tamam/okay*" followed by her final remark that there is no problem with the change and AY's use of the minimal expansion acceptance token "*tamam/okay*" in response to this.

Through a collaborative process of suggestion and decision making, extract 7b demonstrates how the participants interactionally deal with the problem identified in the first phase of the IR interaction with respect to the possibility of "*might*" as a correct answer besides the expected answer "*should*" in a binary-choice item given in a text. In this extract, CI offers two suggestions to replace "*might*" in the options (i.e., *shouldn't*, *don't have to*), which are both negative modal verbs. However, before orienting to the negative versions of the modal verbs, CI first carefully considers the other modal verbs studied in class (i.e., *have to* and *must*). Realizing that they can also be possible as correct answers due to their close association with "*should*", she then looks for other ways of formulating a distractor modal

verb and hence, suggests the negative versions of “*should*” and “*have to*” respectively as a replacement of the trouble source “*might*” in the options. In this way, CI tests all the possible alternative modal verbs studied in class before making a suggestion and deciding on a change. Once she establishes common ground in the group on their assumed unlikelihood as a distractor in and through talk, she moves on to offer suggestions that are open to negotiation and have the possibility of being accepted by the group. As a result, although the group rejects “*shouldn’t*” based on the account that it would make the item too easy and obvious, they eventually decide on the last offered modal verb “*don’t have to*” as an appropriate distractor for the binary-choice item. In this respect, it is seen that one suggestion paves the way for another suggestion in the talk-in-interaction as the participants are constantly in search of a better alternative. When the negative “*should*” is rejected, CI tries the negative version of another modal verb (i.e., *have to*) in the course of deciding on a change and solving the problem in the item.

A collaborative dynamic process is followed via some interactional actions while seeking a change for the problematic option and the suggestions are not simply listed one by one as distinct entities. To begin with, account-giving is one of the predominant features of the interaction in extract 7b, playing an important role in the acceptance of the final suggestion (i.e., *don’t have to*) offered by CI. A thorough account-giving is employed in support of “*don’t have to*” and the suggestion is justified based on evidence. Such an interactional act enables the group to evaluate and discuss the suggested change critically from different perspectives. For instance, CI presents evidence from the surrounding context in the text by translating the key parts of the target sentence into Turkish and referring to the expressions that give a clue about the unlikelihood of “*don’t have to*” within the given context. She also translates the whole sentence into Turkish with “*don’t have to*” inserted into it and especially highlights the verb expressions “*be careful*” and “*protect*” as clues for eliminating “*don’t have to*”, which is also evident in her non-verbal behaviour as she underlines, circles and points at those verbs with her pencil and moves her pencil along them on her paper. By referring to the text and making a connection between the information in the text and the negative modal verb “*don’t have to*”, she shows the meaninglessness of the reason-result relation in the sentence when “*don’t have to*” is inserted into it.

In addition, CI considers her suggestion from the students’ perspectives by explaining that the suggested modal verb would make the students think more about the item without

directly selecting “*should*”, which at the same time displays an orientation to CI’s situated understanding of how a test should be. Likewise, CI evaluates the verb “*protect*”, which is key to the answer, in terms of the students’ knowledge, indicating that it should be within their epistemic domain. This brings a different dimension to the matter at hand and further verifies the appropriateness of the suggested change. Additionally, during her account-giving for her suggestion, CI uses the group’s interactional discussion on “*don’t have to*” as a proof of the thought-provoking aspect of her suggestion (line 83), which makes it more acceptable by AY as she confirms CI and starts to mark the change on her paper. This at the same time makes another orientation to the group’s situated understanding of the norm for testing and thus, to an assumed shared testing principle concerning the level of the item. In other words, it is expected that the item is not too easy or obvious and that the options make the students think carefully about the item.

Even when a mutual understanding is reached about the change, which does not project an expansion, it is seen that the process of account-giving and elaboration of the change made in the item continues in a non-minimal post-expansion sequence (Stivers, 2013). In this post-expansion sequence, CI not only reformulates her prior account for further clarification but also presents further evidence for the validity of the change with an orientation to the verb “*protect*” as the main clue in the text for eliminating “*don’t have to*”. In this way, shared understanding is ensured in the group with respect to the change made in the item. Furthermore, the interactional act of account-giving either for or against a suggestion is employed collaboratively through the participants’ (re)formulations (Heritage & Watson, 1979), AY’s anticipatory completion (Lerner, 1996a) as well as CI’s modified repeat (Kim, 2002; Stivers, 2005). Such a process of account-giving in collaborative turn sequences enables the participants to display agreement and affiliation with a claim, which affects the decision-making process. Based on this, by means of such interactional resources, the participants do not only show agreement but also claim their epistemic right to confirm an assertion, indexing epistemic authority and control over the matter at hand since they all have equal epistemic access to the text in the test material and are in K+ position in terms of their English language proficiency as English teachers (Heritage & Raymond, 2005; Stivers, 2005).

Another recurring interactional action emerging in extract 7b is the challenging of a suggestion or an alternative modal verb, which takes place three times in the talk-in-

interaction until a decision is made on the change. It starts with CI's challenging the other modal verbs studied in class (i.e., *have to* and *must*) in terms of their unacceptability as a distractor in the binary-choice item. Then AY takes the role of checking and challenging the suggestions offered by CI. As a result of the act of challenging a suggestion, the participants either eliminate the offered alternative by backing down and withdrawing it as in the case of "*shouldn't*" or make a counter-challenge in a non-minimal post-expansion sequence by conjointly providing an account for the suggestion as in the case of "*don't have to*" and then reaching an agreement. Thus, projecting an expansion, the act of challenging a suggestion enables the participants to further evaluate the change from different dimensions and makes account-giving relevant next as a counter-challenge to the disagreeing participant. In this respect, the act of challenging a suggestion can be considered to be an important contribution to the validation process since it prompts a critical and careful thought on the suggestion in and through talk (Fulcher & Davidson, 2007). As a result, such a collaborative way of understanding the suggestions leads to the final decision with respect to the change to be made in the problematized item "in open discussion, through a process of dialogue and disagreement" (Fulcher & Davidson, 2007, p.21).

Like in some of the other extracts, it is seen that the suggestions provided in extract 7b are downgraded by CI through their presentation in the form of a confirmation request. CI frames her confirmation requests in such a way that they display an orientation to an assumed shared testing principle concerning the level of the item that it should not be too easy or too difficult but at a moderate level (*çok mu basit olur/would it be too easy, daha zorlar mı acaba?/I wonder if it would make it more difficult*) and hence, exhibit the group's situated understanding of the norm for testing. By downgrading her suggestions with an orientation to an assumed shared testing principle, CI does not implicate epistemic priority as the first provider of a suggestion (Heritage & Raymond, 2005) while the tester AY is also searching for an alternative but invites collaborative inquiry into her suggestions, sharing the responsibility for the decision. This is because (1) the participants both have equal epistemic access to the text in the test material as well as the textbook, (2) they are in K+ position in terms of their English language proficiency as English teachers (Heritage, 2013), and (3) there are certain testing principles that the co-participants in the IR group adhere to in the construction of a language test, sharing a common ground.

Such an approach inviting collaborative inquiry into the suggestions allows for a careful thought on the suggested modal verbs and thus, a ‘healthier’ decision-making with respect to the modal verb that is to replace the distractor “*might*” in the options. Eventually, following the offer of a suggestion, a gap of silence ensues as AY directly starts to think about the suggestions by, for instance, inserting “*don’t have to*” into the target sentence and CI rechecks her suggestions by keeping her gaze on the item. This forms the starting point for account-giving in support of a suggestion and/or challenging it until a decision is made on a modal verb. On the other hand, in response to AY’s display of uncertainty about the last offered suggestion, CI conveys a stance of certainty towards her suggestion by giving thorough account and concrete evidence with a stronger evaluative stance and reasserting her claims, which seems to play an important role in the acceptance of her suggestion by AY. Finally, another interactional resource emerging in the phase of SC in this extract is the use of a preface which projects the upcoming change process for the problematized option. In this extract, CI does it twice, first to initiate a new sequence after the item is problematized; second to indicate the transition back to their search for an alternative after her first suggestion is rejected, which ensures continuity in the suggestion-making sequence.

Regarding the institutional role distribution of the participants in extract 7b, there are indications of tester and teacher roles and responsibilities in this IR interaction. To begin with the tester roles, upon CI’s comment on the unacceptability of “*have to*” and “*must*” as a distractor, the tester AY makes reference to the past item writing process by indicating that “*must*” was an alternative that she thought of while constructing the test but she could not include it as a distractor in the binary-choice item due its close association with “*should*”. This signals her right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item as the writer and thus, the ‘owner’ of the item. Thus, she asserts epistemic priority over the matter by showing that she has already considered all the other modal verbs studied in class as alternatives for “*should*”. In this way, she indirectly states why she had to construct the item in such a way. In addition, AY orients to the textbook as soon as the problem in the item is announced in order to check if there was any other modal verb studied in class, which demonstrates that the textbook is a key source for the testers in preparing syllabus-based language tests and in justifying their choices. Hence, closing the textbook after scanning through it and making reference to an absence of past instructional events (Can Daşkın, 2017), she further proves that there was no other modal verb that escaped her notice during the process of writing the item.

As the tester who is accountable for the test, AY also approaches the suggested changes with caution by carefully thinking about them before displaying her stance. For instance, despite the concrete evidence and account provided by CI, AY plays the devil's advocate and challenges the suggested modal verb "*don't have to*" in line 44 to make sure that it does not fit well in the sentence. Similarly, just as she crosses out "*might*" on her paper to mark the new change, she stops and produces thinking face with her hand on her forehead, carefully thinking about the appropriacy of the change once again. Her display of uncertainty about the change throughout the interaction both verbally and non-verbally prompts further account-giving by CI until mutual agreement is reached. Thus, AY seems to monitor the change process as the tester who has the responsibility to verify the change before showing an uptake of it and finalizing the item. This is the reason why AY also reads aloud the decided modal verb while writing it down on her paper as a replacement of "*might*". It functions as a final check for the suggestion and ensures shared understanding regarding the change made in the item. With respect to the teacher roles in extract 7b, the reviewer CI orients to her institutional role as a teacher by framing her accounts also from the students' perspectives. This implies the teachers' epistemic right, responsibility and obligation (Heritage, 2013) to review the item with consideration of the context of learning and the needs of students and thus, to create the interplay between teaching and testing.

Despite an orientation to the institutional role distribution of the participants, they happen to share the responsibility for the change process in the problematized item. This is most evident in the participants' constant search of an alternative for the problematic option and their careful thought on the offered suggestions. Their production of a thinking face, AY's textbook search, their non-verbal act of mouthing some words, reading some parts of the target sentence silently, inserting the suggestions into the sentence and reading it aloud in Turkish as well as their continuous gaze at the item on their papers (Goodwin, 1981; Rossano, 2013) all illustrate the participants' equal and active engagement in the task and their shared responsibility for the revision of the item. In addition, the participants do not behave assertively but invite collaborative inquiry and give some time to each other to think about a suggestion and to express their stance. As a result, the tester AY, who is relatively more passive in the talk-in-interaction, also makes an important contribution to the interactional process by, for example, challenging a suggestion, offering a (re)formulation, and displaying a stance of agreement or disagreement before making a final decision.

The participants' equal engagement in phase SC can be explained in terms of their assumption of equal epistemic state and authority. They both put themselves in K+ position as reviewers, teachers and testers who have equal epistemic access to the resources and who share a common ground in terms of the institutional expectations with regard to test construction (Heritage, 2013). An orientation is shown to this by CI's presentation of her claim with respect to "have to" and "must" as assumed knowledge and by her request for confirmation from AY for her suggestions with reference to assumed shared testing principle regarding the level of the item. Moreover, CI's use of an inclusive language while projecting a suggestion in the preface (line 23) ascribes an equal epistemic state to the whole group and makes reference to their shared responsibility for the change process in the item and their joint decision-making. Finally, AY's reading aloud while writing down the decided option on her exam paper signals that the change is considered as the product of the whole group.

Extract 8b presents the phase of SC in the reviewing interaction on a multiple-choice reading comprehension item which was problematized in extract 8a. It is to be remembered that in this extract, the teacher AD teaching at B1 plus level does the reviewing of the item and provides her feedback to the tester ME. In the first segment of the extract (i.e., extract 8a, see Section 4.1), AD problematized the verb "underrecognize" in the sentence "moulds have been an underrecognized health problem" given in the part of the text that the item was based on. She claimed that the students would not understand the meaning of "underrecognize" due to their lack of knowledge of the prefix "under", which would then prevent them from finding out the correct answer. Thus, in the following segment of the extract (i.e., extract 8b), having already announced and agreed with the problem, the IR group finds an alternative expression for the verb "underrecognize", which is key to the correct answer through a process of suggestion and decision making.

#### Extract 8b. Moulds, Segment b

- 1 ME: ya da şöyle dicesiz (.) m1- (0.3) mou:lds-  
**or we will say like this**  
 + orients to the reading text on a separate copy and points at the target sentence with a pen  
 + AD looks at the target sentence in the reading text
- 2 AD: have been <neg↑LECTed> diyebiliriz  
**we can say**  
 + shifts gaze to ME
- 3 (0.2)
- 4 [<neglected>

5 ME: [huh  
6 ya ↑da (.) failed (.) to recognize diyebiliriz.  
**or** **we can say**  
+ AD turns gaze back to her paper + shifts gaze to AD  
7 AD: SÜPER olur  
**it would be super**  
+ ME turns gaze back to the reading text  
8 (0.7)  
9 ME: have been failed (.) [<sup>o</sup>to<sup>o</sup>-]  
+ moves the copy of the reading text towards her in an attempt to mark the change on the target sentence  
10 AD: [HAyır ] mould- (.) er olmaz  
**no** **not possible**  
11 ((AD bends forward, takes a closer look at the target sentence and scans through it for 2.0s while ME sustains her gaze at the sentence))  
12 ME: people ↑have  
13 AD: <when> (.) [er er er hh (.) ↑SCI ]ENTists  
14 ME: [failed to recognize health-]  
15 AD: diyelim  
**let's say**  
16 <sup>1#</sup>(1.6)  
17 scientists o o  
**that that**  
18 (0.4)  
19 health care <sup>#1</sup> professionals'la o↑lur  
**would be okay with "health care professionals"**  
+ shifts gaze to ME and then nods head  
#1 (lines 16-19) ME writes "scientists" in place of "moulds" in the target sentence in the text  
20 ME: evet (.) scien[tists  
**yes**  
+ AD turns gaze back to her paper  
21 AD: [scientists <sup>2#</sup>[↑have]  
22 ME: [have ] failed  
23 AD: failed (.) to  
24 (1.0)  
25 <recognize>  
26 (2.2) <sup>#2</sup>  
#2 (lines 21-26) ME writes the new verb expression "have failed to recognize" on her copy of the reading text as a replacement of "have been an underrecognized" given in the text while ME and AD read it aloud through collaborative turn completions  
27 health problem demiyelim buna artık  
**let's not say "health problem" to this anymore**



28 (0.3)  
 29 ME: [recogni:ze ]  
 + reads out from the sentence she has just noted on the reading text  
 30 AD: [er mu moul-] (.) MOULD (.) <cases>  
 + shifts gaze to ME and then back to her paper  
 31 ME: evet  
**yes**  
 + positions hand on the reading text in an attempt to make the next change in the target sentence  
 32 ((ME stops for a while, produces thinking face and then writes "mould cases" as a replacement of "health problem" in the target sentence in the text while AD keeps her gaze on her paper))  
 33 AD: tamam mı?=  
**is that okay?**  
 + nods head  
 34 ME: =<sup>o</sup> huh-huh<sup>o</sup> [but the situation is changing]  
 + reads aloud from the target sentence in the text  
 35 AD: [böyle daha güzel oldu ]  
**it is better like this**  
 36 ME: evet (.) [aynen ]  
**yes exactly**  
 37 AD: [böyle daha] güzel oldu.  
**it is better like this**  
 38 (1.0)  
 39 ta↑mam  
**okay**  
 40 ME: tamam  
**okay**

Extract 8b starts with the tester ME's initiation of a new sequence through her use of a preface with an inclusive language (*dicez/we will say*) which projects the announcement of an upcoming suggestion for the problematic verb "underrecognize" given in the target sentence in the reading text. ME's preface is accompanied by her orientation to the reading text on a separate copy and pointing at the target sentence with a pen while the reviewer AD is looking at it on her exam paper. ME then reads aloud the subject of the sentence (i.e., *mou:lds*) with an elongation following an intra-turn pause and a cut-off in an attempt to produce her suggestion for the verb "underrecognized" that comes next in the sentence. AD then takes over to complete ME's turn with her suggestion "<neg↑LECTed>" in line 2 using an inclusive language with a shift of gaze to ME. Thus, AD completes the reading aloud of the first part of the sentence that was initiated by ME but

she does it by embedding her suggestion in the verb expression in the original sentence (i.e., have been <neg↑LEcted>). AD articulates the verb in a slower pace by putting an emphasis on it and increasing the volume and pitch of the second syllable of the verb. AD then repeats the verb she has just suggested in an isolated form (i.e., <neglected>) in line 4 in a slower pace and with an emphasis once again.

In overlap with AD's turn, ME displays agreement through an acknowledgement token but then just as AD turns her gaze back to her paper, using an inclusive language, ME offers the second suggestion "failed (.) to recognize" as a rephrasing of "under" in the verb "underrecognized" in line 6 with a shift of gaze to AD. This is quickly met with an affiliative responsive uptake by AD, who accepts the new suggestion through an upgraded positive assessment of it (SÜPER olur/*it would be super*) in line 7 whereby the positive evaluation token "SÜPER/*super*" is produced at a higher volume. In this way, AD displays a stance of agreement and affiliation with ME's suggestion.

Turning her gaze back to the reading text, ME in line 9 reads aloud the target sentence starting with the present perfect passive form (i.e., have been) by inserting the verb expression she has just suggested in place of the trouble source "underrecognized" with an emphasis on "failed". However, before she completes her reading out of the rest of the sentence and just as she moves the copy of the reading text towards her in an attempt to mark the change on the target sentence, AD overlaps in line 10 and cuts off ME's utterance. Identifying a problem with respect to the accuracy of the sentence when "underrecognized" is directly substituted with the suggested rephrasing, AD initiates a correction sequence. Due to the mismatch between the new verb expression and the subject of the sentence "moulds", AD uses the correction preface "Hayır/no" with a louder volume in word-initial position and then refers to the trouble source "moulds" in the reformulated target sentence by stating its inaccuracy as the subject of the suggested verb expression. As a result, AD bends forward and takes a closer look at the target sentence, scanning through it for a while in silence as ME sustains her gaze at the sentence in search of an alternative for the trouble source "moulds" in the subject position of the sentence.

In line 12, ME takes over to perform a self-correction by suggesting "people" for "moulds". In doing this, she uses an embedded correction as she reads aloud the target

sentence with her suggestion “people” embedded into it along with the transformation of the passive structure (have been failed) into its active form (↑have failed). In overlap with ME’s turn, following her hesitation markers as she is in search of an alternative with her gaze at the sentence, AD proposes “↑SCIENtists” as another correction for “moulds” in line 13 with an initial rise in the pitch and the volume of the word as well as a word-final emphasis. She also uses an inclusive expression of suggestion in line 15 (diyelim/let’s say). AD’s suggestion cuts off ME’s utterance as she cannot complete her reading out of the final word of the sentence. This is received with ME’s non-verbal agreement and uptake as she directly writes “scientists” as a replacement of “moulds” in the target sentence. In this way, the sentence becomes subject to a further reformulation after deciding on a change in the previously problematized verb expression of the sentence. Meanwhile, AD provides an account for her suggestion in line 19 by explaining that “scientists” conforms with the other expression (i.e., health care professionals) used in the paragraph to refer to people who study moulds. Through the end of her turn-at-talk, she shifts her gaze to ME, nodding her head and thus, further acknowledges her suggestion “scientists”.

In line 20, ME explicitly shows agreement with and a stance of acceptance of AD’s suggestion through her use of the compliance token “evet/yes” followed by her reading aloud of the reformulated sentence starting with the suggested correction “scientists” already written on ME’s text. At this point, AD turns her gaze back to her paper and then she also starts to reproduce the new sentence. Then between lines 21 and 25, a collaborative writing activity takes place as ME writes the newly formulated verb expression “have failed to recognize” on her copy of the reading text as a replacement of “have been underrecognized” in the original sentence while ME and AD reproduce it through collaborative turn completions. This further demonstrates the group’s agreement and final decision on the change and ensures shared understanding in the group with respect to the change made in the sentence which has undergone several changes.

However, AD stops reading aloud the sentence when she comes to its object. Following a gap of silence when ME continues to make the change on her paper, AD this time orients to the object of the sentence (i.e., health problem) in line 27 and initiates correction with respect to it by suggesting to change this expression in the object position with an inclusive

language. After the substitution of “*moulds*” with “*scientists*”, “*health problem*” is now inappropriate as the object of the sentence since it provides missing information on its own while the paragraph talks about the health problem caused by moulds in particular. Thus, the issue of moulds needs to be mentioned somewhere in the sentence. As a result, both ME and AD start to think of a suggestion for “*health problem*”. This is evident in ME’s reading aloud the verb “*recogni : ze*” just preceding the object of the sentence with an elongation in line 29 and AD’s use of a hesitation marker as well as an intra-turn cut-off in line 30.

Then in line 30, AD offers the suggestion “MOULD (.) <cases>” in place of “*health problem*” by shifting her gaze to ME and then back to her paper. She produces “MOULD” at a higher volume with an emphasis and articulates “<cases>” in a slower pace. This is received with ME’s display of agreement through the use of the compliance token “*evet/yes*” in line 31 just as she positions her hand on the reading text in an attempt to make the next change in the target sentence. However, before writing “*mould cases*” as suggested by AD to change “*health problem*”, ME stops for a while and produces thinking face. Only then does she write “*mould cases*” as a replacement of the object “*health problem*” in the sentence while AD keeps her gaze on her paper. As a result, the target sentence in the text undergoes another reformulation after the acceptance of ME’s suggestion for the previously problematized verb “*underrecognize*”.

In line 33, AD solicits ME’s confirmation with respect to the finalized sentence accompanied by her head nod. This is immediately received with an acknowledgement token from ME in line 34, who then reads aloud the *but*-clause following the object of the target sentence and does a final check. Simultaneously with ME’s reading aloud, AD offers a positive assessment of the change in line 35 and repeats it in line 37. ME further confirms AD and displays agreement through her use of two compliance tokens successively (i.e., *evet/yes*, *aynen/exactly*) in line 36. Following a gap of silence, AD requests ME’s final confirmation with respect to her acceptance of the change and proposes sequence closure by uttering the acceptance token “*ta ↑ mam/okay*” with a rising intonation. This receives a confirming response from ME and thus, the sequence closes once ME’s acceptance of the change is ensured.

Extract 8b demonstrates the suggestion and decision making process to change the previously problematized verb “*underrecognize*” in the sentence “*moulds have been an underrecognized health problem*” given in the part of the text that the multiple-choice reading comprehension item was based on. In order to replace this verb with an expression that the students can understand for them to be able to answer the item as intended, two suggestions are offered (i.e., *neglected*, *failed to recognize*) and the group decides on the second one (i.e., *failed to recognize*). However, different from the other extracts, the change process continues in extract 8b since the suggested change for the verb of the target sentence requires other changes and corrections in the subject and object of the sentence for the new verb to fit meaningfully and accurately in the sentence. As a result, another suggestion and decision making process starts for the changes in the subject and object of the sentence. In the end, it is seen that with the change of the problematic verb “*underrecognize*”, the target sentence undergoes several changes and reformulations until it is finalized.

Different from extracts 1b, 4b, and 5b, in this extract, it is not the suggestion for change itself that is collaboratively reformulated but the location of the trouble source (i.e., the target sentence in the text) that is reformulated for it to be meaningful. Since the suggestion “*failed to recognize*” fulfils the group’s expectations and is, thus, met with their uptake, rather than seeking other alternative verb expressions, they together work towards making it fit meaningfully in the sentence. Hence, the original sentence “*moulds have been an underrecognized health problem*” is eventually transformed into “*scientists have failed to recognize mould cases*” where the subject and the object of the sentence as well as the form of the sentence from passive to active had to be changed along with the change of the verb expression. In this way, the students are expected to make a better sense of the sentence and find out the correct answer without being punished for their lack of knowledge of the prefix “*under*” used in such verbs as they have not studied those prefixes yet. Hence, extract 8b shows that when a change is made in the test, the participants need to be cautious about it as it might require other change and correction processes. In other words, they should be alert to any other problems that might emerge after a particular change is made in the test material.

There are certain interactional resources that play a key role in prompting a collaborative change process in extract 8b. To begin with, ME’s preface makes the suggestion-making process relevant next, enabling the participants to immediately seek a suggestion for the

problematic verb in the sentence. In addition, ME's reading aloud of the target sentence and its collaborative completions by AD is a predominant interactional resource, which serves several functions in the talk-in-interaction in extract 8b. Firstly, it helps the group to think of a suggestion for change, considering the target sentence as a whole and paves the way for its announcement. For instance, ME reads out the verb "*recognize*" in the newly formulated sentence with an elongation in line 29 while thinking of an alternative for the object of the sentence that just follows it. Similarly, AD offers the first suggestion by completing the reading aloud of the sentence that was initiated by ME in line 1. ME reads out the subject of the sentence (i.e., *moulds*) with an elongation, which projects the continuity of her turn with a suggestion for the verb of the sentence. In this example, the reading aloud not only guides ME in searching for a suggestion but also allows AD to offer her suggestion by completing ME's turn of reading aloud in an alignment, which demonstrates their shared purpose.

Secondly, the act of reading aloud enables the group to check the appropriacy of a suggestion in the target sentence. For example, ME utters her suggestion "*people*" and at the same time checks its appropriacy by reading aloud the target sentence with "*people*" directly inserted in place of "*moulds*" in the form of an embedded correction and the passive structure changed into its active form. ME does the same in line 9 where she inserts her suggestion "*failed to recognize*" in the present perfect passive form given in the target sentence. This acts as a check for the group to see if it fits well in the sentence before marking the change on her paper. Likewise, ME checks the appropriacy of the suggestion "*mould cases*" offered for the object of the sentence by reading aloud the rest of the sentence (i.e., the *but*-clause) that follows it after noting the change on her paper. In this way, she examines if the change conforms with the whole sentence and causes no inconsistencies.

Thirdly, the act of reading aloud leads to the realization of a problem in the sentence and thus, the initiation of a correction sequence. To illustrate, ME's reading aloud of her suggestion by embedding it in the present perfect passive form given in the target sentence enables AD to identify the mismatch between the suggested verb "*failed to recognize*" and the subject of the sentence "*moulds*". Likewise, ME and AD's collaborative writing activity allows AD to find out that the object of the sentence is also inappropriate after the changes made in the verb and the subject of the sentence. This leads to the initiation of another correction sequence and eventually the careful reformulation of the target sentence to solve the identified problem.

As in most of the extracts, in this extract too, account-giving for a suggestion is employed in the talk-in-interaction. AD does it in line 19 where she provides an account for her suggestion “*scientists*” offered for the subject of the sentence “*moulds*”. In this way, right after making the suggestion, AD carefully thinks about her suggestion and makes an evaluation of it in terms of its appropriacy within the context given in the text and thus, its relevance to testing principles. Selection of such a word as “*scientists*” which corresponds to the expression “*health care professionals*” in the text would allow the test-takers to understand the relation between parts of a text through lexical devices and contribute to the coherence of the text (Heaton, 1990). As a result of AD’s account-giving, mutual understanding is reached about the suggestion.

Throughout the correction sequences in extract 8b, AD adopts a stance of certainty with her suggestions for the subject and object of the target sentence that are to be revised after the group decides on a change to overcome the problem in the sentence. This is evident in her declarative statements while providing her suggestions and her account as well as her articulation of her suggestions with a higher volume and an emphasis. Therefore, her suggestions are directly accepted without a need for the group to seek other alternatives. Furthermore, the participants convey their stance of acceptance of a suggestion by means of different interactional resources such as (1) offering a positive assessment as AD does in lines 7, 35, and 37, (2) non-verbally noting the suggested change on one’s paper as ME does in the cases of the suggestions “*scientists*” and “*mould cases*”, and (3) employing a collaborative writing activity. By means of such interactional resources, the participants show their understanding of and affiliative stance towards a change, which in turn affects the decision-making process in the interaction.

Considering the institutional role distribution, extract 8b seems to display an orientation to the tester roles and responsibilities. It is most evident in ME’s making the changes on her copy of the reading text as the tester who has the responsibility and obligation to finalize the test and make it ready for administration. Additionally, as the tester who is accountable for the test, ME approaches the suggested changes with caution and monitors the change process by carefully thinking about them. She especially does it by inserting the suggested changes into the target sentence and then reading it aloud with the embedded changes and corrections. Her reading aloud functions as a final check for the suggestion, which leads to either the initiation of a correction sequence or its acceptance, and solicits mutual

understanding regarding the changes in the item. Similarly, upon AD's suggestion for the object of the sentence, although she positions her hand on the reading text in an attempt to write the suggested word "*mould cases*" in the target sentence, she stops for a while and produces thinking face and only after verifying it, she makes the change.

Despite an orientation to the institutional role distribution of the participants, they happen to share the responsibility for the change process in the problematic item. Both the tester and the reviewer actively seek alternatives for the verb of the sentence, which was previously problematized. ME, for instance, does not directly orient to AD's suggestion but offers another alternative verb expression in search of a better one, leaving it open to negotiation. As a result, the group decides on her suggestion. A similar process is followed in the case of the subject of the sentence where AD does not orient to ME's suggestion (i.e., *people*) but offers a new suggestion (i.e., *scientists*) that is met with the group's uptake. As a result, a collaborative suggestion and decision making process takes place in the talk-in-interaction in extract 8b. In addition, the participants' online thinking of a suggestion for the located trouble sources in the target sentence is evident in their intra-turn pauses and cut-offs, hesitation markers, elongated reading aloud of a word, non-verbal act of scanning through the sentence as well as their continuous gaze at the sentence on their papers (Goodwin, 1981; Rossano, 2013). All these interactional resources illustrate the participants' equal and active engagement in the task and their shared responsibility for the change in the item.

The participants' use of an inclusive language (e.g., *diyebiliriz/we can say*, *diyelim/let's say*), their collaborative writing activity and AD's final confirmation checks with respect to ME's acceptance of the changes made in the item are the other indicators of their shared responsibility for the change process in the item and their joint decision-making, illustrating their ownership of the test as the shared product. This can also be explained in terms of the participants' assumption of equal epistemic state and authority since they both have equal epistemic access to the reading text and are in K+ position in terms of their English language proficiency as English teachers (Heritage, 2013). In short, this extract indicates that the participants in the IR interaction are actively processing as they are suggesting and deciding on a change to improve the item.



Extract 9b presents the phase of SC in the reviewing interaction on a multiple-choice reading comprehension item which was problematized in extract 9a. It is to be remembered that in this extract, the two teachers GE and DU teaching at B1 level do the reviewing of the item and provide their feedback to the tester SA. In the first segment of the extract (i.e., extract 9a, see Section 4.1), GE problematized the distractor option “d” in the multiple-choice reading comprehension item which asks about the opinion of the famous striker Michael Owen’s father with regard to his son’s profession and his popularity. GE claimed that option “d”, which states “*people can freely criticize Michael when he does not play*” could also be possible as a correct answer when the students interpret the meaning of “*can*” given in the option as a possibility rather than a permission and that the item might end up having two possible correct answers. Thus, in the following segment of the extract (i.e., extract 9b), having already announced and agreed with the problem, the IR group revises option “d” through a process of suggestion and decision making for change and improvement in the item.

#### Extract 9b. Freely Criticize, Segment b

- 1 SA: o zaman onu nasıl yapabiliriz ki?  
**then how can we make it**  
+ all look at their exam papers
- 2 (2.3)
- 3 GE: ↑olumsuz yapabilirsin  
**you can make it negative**  
(0.2)
- 4 (0.2)
- 5 SA: should desek?  
**what if we say "should"**  
+ shifts gaze to DU  
+ DU and GE shift gazes to the item on the next page
- 6 (0.7)
- 7 GE: should [criticize. ]
- 8 SA: [can yeri↑ne (.) çünkü] aslında babası  
+ shifts gaze to GE + shifts gaze to DU
- 9 >insanlar criticize< etsin demi↑yo  
**instead of "can" because his father doesn't actually  
say people shall criticize**  
+ DU shifts gaze to SA  
+ GE shifts gaze to the relevant part of the  
text on the previous page
- 10 (0.3)

- 11 DU: de[miyo]  
**doesn't say**  
+ turns gaze back to her paper, nodding head
- 12 SA: [kötü] bişey diyo  
**he says it is something bad**  
+ shifts gaze to GE
- 13 GE: evet  
**yes**
- 14 DU: evet  
**yes**  
+ nods head
- 15 SA: o zaman should koyduğumuz za↑man (.) yanlış oluyo sık  
**then when we put "should", the option becomes wrong**  
+ GE turns gaze back to the item on the next page  
+ turns gaze to her paper and then to DU
- 16 (0.2)
- 17 DU: should criti[↑ci:ze
- 18 GE: [shou[ld evet  
**yes**  
+ shifts gaze to SA, nodding head
- 19 SA: [huh-huh
- 20 (0.2)
- 21 DU: >michael< when he does[n't [play =olur  
**it is okay**  
+ nods head
- 22 SA: [di mi? [>aslında <sup>1#</sup>babası<  
**doesn't it? actually his father**  
+ turns gaze to her paper
- 23 GE: [tamam  
**okay**  
+ turns gaze back to the item
- 24 SA: [çün↑kü (.) onu eleştiriyo yani herkes er  
**because he criticizes that, that is to say everyone**  
+ DU shifts gaze to SA's paper and then back to hers
- 25 GE: [pekala  
**alright**
- 26 SA: <sup>2#</sup>çocuğu↑mu <sup>#1</sup>  
#1 (lines 22-26) SA crosses out "can" in option "d" and writes "should" in its place
- 27 (0.4)
- 28 [sert <ediyo> diye kötü- disadvantage diyo: ]  
**does harsh to my kid, it is bad- it is a disadvantage he says**  
+ shifts gaze to GE and then to DU

29 DU: [eleştiriyo ama ↑bu (.) ünlü olmanın bir şeyi]  
**he criticizes but this is one thing about being famous**  
+ shifts gaze to SA

30 (0.2)<sup>#2</sup>  
#2 (lines 26-30) GE nods head

31 [hani  
**well**

32 GE: [getirisi[dir [falan diyo evet  
**its result or so he says yes**

33 SA: [eVET  
**yes**

34 DU: [dezavantajı diyo  
**its disadvantage he says**  
+ nods head + shifts gaze to GE, then to her paper

35 SA: [<getirisi> ama dezavanta[jı diyo]  
**its result but its disadvantage he says**  
+ turns gaze to her paper and points right hand to the text

36 DU: [o zaman]

37 tamam [olur ]  
**okay then it is possible**

38 SA: [di mi?] should dediği[miz zaman ]  
**doesn't he? when we say "should"**  
+ shifts gaze to GE

39 DU: [ ° huh-huh ° ]  
+ nods head

40 GE: evet evet=  
**yes yes**

41 SA: =en azından öyle bi öneri (.) vermesi söz konusu  
+ GE turns over the page

42 değil  
**at least him giving such a suggestion is not the the point in question**

43 DU: de[ğil  
**not**

44 GE: [değil  
**not**

45 (0.5)

46 DU: doğ[ru  
**right**  
+ SA turns gaze back to her paper

47 SA: [tamam  
**okay**

As all the participants are looking at their exam papers, extract 9b starts with the tester SA's preface with an inclusive language which projects the upcoming suggestion and decision making process for the change in the problematic option "d". Following a gap of silence as the participants are thinking of a solution, GE makes the first suggestion in line 3 with a turn-initial pitch rise. She suggests making the option negative to make it incorrect as a distractor. Without orienting to GE's suggestion, using an inclusive language, SA offers a different alternative in lines 5 and 8, that is replacing "*can*" in the option with the modal verb "*should*". She directs the suggestion to the other participants with a shift of gaze first to DU and then to GE and a slight rising intonation. Her suggestion draws the other participants' attention as they both turn their gazes to the item on the next page to check its appropriacy. In doing this, GE, in line 7, reads aloud the verb expression in the option by inserting the suggested modal verb "*should*" into it (i.e., *should criticize*).

SA proceeds with her turn between lines 8 and 12 by providing account for her suggestion with a shift of gaze to DU and GE respectively. By making reference to the relevant context in the reading text, SA explains that the famous striker Michael Owen's father is not actually in favour of the criticisms levelled at his son but rather claims that it is something bad. This presents evidence for the incorrectness of option "d" as a distractor when "*can*" is replaced with "*should*". SA articulates the negative verb "*demi ↑yo/doesn't say*" by putting an emphasis on the first syllable and increasing the pitch of the final syllable, which highlights the father's negative remark with regard to the criticisms directed at his son. As SA is giving her account, DU shifts her gaze to SA, maintaining mutual gaze with SA for a short time and GE turns her gaze to the relevant part of the text on the previous page as referred to by SA.

At the transition-relevance place upon SA's first possible completion of her unit of talk, DU uses a partial modified repeat (Kim, 2002; Stivers, 2005) of SA's account in line 11 by repeating the final word of SA's turn (*demiyo/doesn't say*), which is the core of her account and thus, confirms SA, displaying an affiliative stance with her. Her turn is accompanied by her head nod as she shifts her gaze back to her exam paper. In line 13, GE also shows agreement through her use of the compliance token "*evet/yes*" and DU repeats the same compliance token in line 14 accompanied by her head nod, which further displays her agreement with SA's account for her suggestion.

Turning her gaze to her exam paper and then quickly to DU, SA orients to the previously announced problem in line 15 by stating that the option becomes wrong when “*should*” is inserted into it, which is the desired outcome. In this way, she solicits the other participants’ confirmation about the change and ensures mutual understanding. She uses an inclusive language when referring to the action of replacing “*can*” with “*should*” (*should koyduğumuz za↑man/when we put “should”*), which is an indicator of their shared responsibility for the decided change in the item. Meanwhile, GE turns her gaze back to the item on the next page. Looking at her exam paper, DU then rechecks the suggestion by embedding it in the verb expression in the option (*should criti↑ci:ze*) and reading it aloud with a rising intonation and an elongation in line 17. Overlapping her at a turn-final position, GE repeats the suggested modal verb “*should*” in line 18 and then quickly acknowledges SA’s remark about the change by using the compliance token “*evet/yes*” with a shift of gaze to SA accompanied by her head nod. This is met with SA’s acknowledgement in line 19. DU continues to read aloud the rest of the option in line 21 accompanied by her head nod as she is rechecking the suggestion within the context given in the option and then immediately confirms and agrees with SA’s suggestion by using the acceptance token “*olur/okay*”. Overlapping DU’s turn, SA does a confirmation check (*di mi?/doesn’t it?*) with respect to the incorrectness of the distractor when “*should*” is used by turning her gaze away from DU to her exam paper.

Overlapping DU’s reading aloud of the option at a turn-final position, SA initiates a non-minimal post-expansion sequence (Stivers, 2013) in line 22 upon the group’s agreement and decision on the change. In this post-expansion sequence, she provides further clarification of her account between lines 22 and 28. Her account-giving is embodied by her crossing out “*can*” in option “d” and writing “*should*” in its place, which demonstrates their ultimate decision on the change. In addition, in her account, SA once again refers to the part of the reading text where the father’s opinion about his son’s popularity is explained and thus, reasserts that the father actually criticizes the public’s harsh reactions to his son and considers it as a disadvantage. Just as she initiates her clarification of her account, GE displays a stance of acceptance of SA’s suggestion through her use of the minimal acceptance tokens “*tamam/okay*” and “*pekala/alright*” in lines 23 and 25 by turning her gaze back to the item. During SA’s turn-in-progress, DU shifts her gaze to SA’s exam paper as SA is marking the change on her paper and then back to hers. Just as SA

finishes making the change on her paper, she continues her account in line 28 by shifting her gaze to GE and then quickly to DU.

At this point, DU overlaps SA's ongoing turn of account in line 29 and provides a complementary explanation by referring to a different part of the relevant context in the reading text, that is the issue of being famous. She explains that the father criticizes bad public reactions, confirming SA but it is mentioned as one of the disadvantages of being famous. Through the end of her unit of talk in line 29, DU shifts her gaze to SA. Moreover, SA and DU's collaborative account between lines 28 and 30 is accompanied by GE's head nod, which displays her agreement with the other participants' explanations. However, before DU completes her account in line 29, she pauses and uses the discourse marker "hani/*well*" as she is searching for a word to express what the father says about the bad public reactions in relation to being famous. As a result, GE employs anticipatory completion (Lerner, 1996a) in line 32 by stating that the bad public reactions are the result of being famous (*getirisidir/its result*). DU does not orient to GE's anticipatory completion but in overlap with GE's turn, she completes her prior turn of account-giving in line 34 by explaining that such public reactions are mentioned as the disadvantage of being famous, which is accompanied by her head nod. Through the end of her unit of talk, she shifts her gaze to GE and then back to her paper.

Meanwhile, maintaining mutual gaze with DU, SA acknowledges DU's explanation through her use of the compliance token "eVET/*yes*" produced at a higher volume in line 33. She then orients to both GE and DU's turns and acknowledges them in line 35 by (re)formulating and summing up their point that the bad public reaction is the result of being famous but it is mentioned as a disadvantage, which is embodied by SA's turning her gaze to her paper and pointing her right hand to the text. In this way, she proves that the father has a negative stance toward the criticisms levelled at his son and that the new change makes the option wrong as a distractor since the use of "*should*" presents the public criticisms as something that is supported by the father.

Upon the participants' collaborative account-giving with reference to the information in the reading text, DU once again displays her stance of acceptance of the change through her use of the acceptance tokens "tamam/*okay*" and "olur/*it is possible*" in line 37.

Overlapping DU's turn at a turn-final position, SA does a final confirmation check (*dimi?/doesn't he?*) and shifting her gaze to GE, she further comments on the irrelevance of the newly changed option to the context given in the reading text due to the meaning "*should*" brings to the option since the father's giving such a suggestion is not the point in question in the reading text. She uses an inclusive language when referring to the act of putting "*should*" in the option. SA's confirmation check is met with both DU and GE's acknowledgements in lines 39 and 40. In addition, DU and GE both use a partial modified repeat (Kim, 2002; Stivers, 2005) of SA's final remark by repeating the last word of SA's turn (*değil/not*), which confirms SA and displays an affiliative stance with her. DU also uses the stance marker "*doğru/right*" in line 46, further acknowledging SA's final comment about the change. As a result, the appropriacy of the change made in option "d" is ensured and mutual understanding is established in the group. Turning her gaze away from GE to her paper, SA then proposes sequence closure with her use of the minimal acceptance token "*tamam/okay*".

Extract 9b demonstrates the suggestion and decision making process to change the previously problematized distractor option "d" in a multiple-choice reading comprehension item. In order to make this option incorrect and unlikely as a distractor, two suggestions are offered successively, which are (1) making the option negative and (2) replacing the modal verb "*can*" in the option with "*should*". The group directly focuses on the second alternative and eventually, decides to put "*should*" in the option after carefully thinking about it. It is seen that even when a suggestion is made, the participants continue to seek other alternatives in search of a better one. For instance, when GE provides the first suggestion of making the option negative, the group does not orient to it but keeps their gazes at the test material. The brief gap of silence that follows with no uptake from the other participants can be interpreted as a display of their dispreference with GE's suggestion.

An interactional action that plays an important role in the acceptance of SA's suggestion is account-giving. Even though SA downgrades her suggestion by directing it to the other participants with a shift of gaze to them and a slight rising intonation, without waiting for their responses, she quickly provides an account for her suggestion by presenting evidence from the reading text and displays a stance of certainty with regard to the appropriacy of her suggestion, which seems to be an important step for the acceptance of her suggestion.

Although the participants' display of agreement with the suggestion does not project an expansion, she even initiates a non-minimal post-expansion sequence where she provides further clarification of her account by making reference to the reading text to prove the appropriateness of the reformulated option. In this way, she ensures that the new option does not pose any risk by turning out to be possible as a correct answer. Thus, as in the other extracts, the suggestion that is justified based on evidence is likely to be accepted and put to use by the group.

With respect to the act of account-giving for a suggestion, which seems to be a predominant feature of the interaction in extract 9b, SA considers her suggestion from different perspectives to prove that the new distractor is not possible as a correct answer and to establish mutual understanding in the group. Firstly, she refers to the information in the reading text and explains the negative comments of Michael's father with regard to the public criticisms directed at his son. This presents contradictory evidence for the correctness of option "d" and validates it as a distractor since the father does not support such criticisms as expressed in the option but instead considers them as one disadvantage of being famous. SA also considers the reading text with reference to the new option in line 41 by making the remark that the father's suggesting the public to criticize his son as claimed in the new option is not the point in question in the text. In this way, she evaluates the change from different dimensions, giving the option a careful and a critical thought. Secondly, SA orients to the violated testing principle in line 15 and states that "*should*" makes the option wrong, which is the desired outcome and the main goal of the change process in extract 9b. In this way, she shows that such a change helps overcome the identified problem in option "d".

Although SA plays the main role in providing a justification for her suggestion, the whole group evaluates the change by, for instance, inserting the suggestion "*should*" in the option and employs the interactional act of account-giving collaboratively. They do this through their supportive and complementary explanations with reference to a different aspect of the relevant context (i.e., the issue of being famous), GE's anticipatory completion (Lerner, 1996a) and their modified repeats of SA's turns (Stivers, 2005). As a result of such a process of account-giving in collaborative turn sequences, the change in option "d" is discussed thoroughly and is eventually taken up by the group. Those interactional resources also enable the participants to display agreement and affiliation with a claim, which affects the decision-making process. Based on this, by means of such interactional resources, the



participants do not only show agreement but also claim their epistemic right to confirm an assertion, indexing epistemic authority and control over the matter at hand since they all have equal epistemic access to the reading text in the test material (Heritage & Raymond, 2005; Stivers, 2005).

Regarding the institutional role distribution in extract 9b, there are indications of tester roles and responsibilities in this IR interaction. Throughout the talk-in-interaction in extract 9b, SA seeks the other participants' confirmation for her claims and thus, for her suggestion and further checks their confirmations, orienting to her role as the tester who is accountable for the test. She also sums up the other participants' explanations through a (re)formulation in line 35 and makes a final remark about the change in line 41 before closing the sequence. In this way, she makes sure that the change she has suggested is appropriate, causing no problem in the item and establishes mutual understanding in the group. For the same reason, although the group's agreement with the suggestion does not project expansion, SA initiates a non-minimal post-expansion sequence where she provides further elaboration of her account for her suggestion and receives the other participants' acknowledgements. This is important for the tester due to accountability issues in testing because in any case of criticism, as the tester of the institution, she will have to explain to the stakeholders why the item is designed in such a way. Furthermore, SA makes the change on her copy of the exam paper as the tester who has the responsibility and obligation to finalize the test and make it ready for administration. Finally, as the tester, SA monitors the change process by means of interactional resources such as (1) applying a preface that projects the upcoming suggestion-making process for the problematic option, (2) initiating a post-expansion sequence, (3) launching sequences of account-giving for her suggestion, (3) closing the sequences by, for instance, summing up the other participants' points or initiating new sequences through her confirmation checks (i.e., *di mi?/right?*), and (4) proposing the closure of the phase of SC through her minimal acceptance token.

As in the case of other extracts, despite an orientation to the institutional role distribution of the participants, they happen to share the responsibility for the change process in the problematic item. To begin with, the participants all seek alternatives to make option "d" incorrect as a distractor and collaboratively provide account for the decided change, which illustrates their ownership of the test as the shared product. SA, for instance, does not orient to GE's suggestion but offers another alternative in search of a better one, leaving it open to

negotiation. As a result, they all get engaged in evaluating and checking the appropriacy of the suggestion by inserting “*should*” in the option and reading it aloud. Based on this, SA’s confirmation requests lead to joint decision-making in the IR group. In addition, the engagement of all the participants in the change process is evident in their gaze behaviour. It is seen that throughout the interaction, the participants continuously look at their exam papers only sometimes directing their gazes at each other for a short time and shift their gazes between the reading text and the item on the next page. However, this does not indicate disengagement; on the contrary, it signals that the participants actively think about the suggested alternative for the problematic option in such a competing activity (Goodwin, 1981; Rossano, 2013). Finally, SA’s use of an inclusive language for three times throughout the interaction with regard to the change to be made in the option once again makes a reference to the participants’ shared responsibility for the change process in the item. All these can be explained in terms of the participants’ assumption of equal epistemic state and authority since they all have equal epistemic access to the reading text and can offer or evaluate a suggestion based on the information in the text.

Extract 10b presents the phase of SC in the reviewing interaction on a matching item which was problematized in extract 10a. It is to be remembered that in this extract, the teacher YU teaching at B1 plus level does the reviewing of the item and provides his feedback to the tester ME. This item appears in a matching test task type designed for a story quiz, which tests students’ comprehension of a story titled “Lorna Doone”. In this matching test, the students are expected to match the descriptions of four characters with one of the six character names given in a box by writing the name of the character in the blanks provided in front of the statements. In the first segment of the extract (i.e., extract 10a, see Section 4.1), YU problematized the description of one particular character named Lizzie, claiming that Lizzie was one of the minor characters in the story whom he had not focused on while reading the story and thus, he was unable to match the statement with the expected answer “Lizzie”. Thus, in the following segment of the extract (i.e., extract 10b), having already announced and agreed with the problem, the IR group finds an alternative character name from the story and formulates a description for it through a suggestion and decision making process. This extract was also presented at a data session. The analysis provided here is, thus, supported by the analysis carried out at this data session.

**Extract 10b. Lizzie, Segment b**

- 1 ME: ↑belki bire başka bişey  
*maybe for one something else*  
+ shifts gaze to her exam paper
- 2 (0.3)
- 3 YU: evet ya >me[sela< ben-  
*well yes for example I-*  
+ looks at his exam paper
- 4 ME: [kimi düşü(.)nürsün=  
*who would you think of*  
+ shifts gaze to YU's paper
- 5 YU: =jeremy stickles (.) hani (.) er: (0.2) is a s↑p:y  
*well*  
+ underlines the name "jeremy stickles" given in the box
- 6 (0.6)
- 7 who helped
- 8 (0.5)
- 9 er:[:
- 10 ME: [↑hnm:  
+ turns to her paper
- 11 (0.8)
- 12 YU: who helped  
+ produces thinking face  
+ ME starts to write down the new sentence and then stops
- 13 (0.9)
- 14 john
- 15 (0.3)
- 16 ME: >°who° ya da< who worked (.) <for the sake of>  
*or*  
+ produces thinking face  
+ YU shifts gaze to ME
- 17 (0.6)
- 18 er: yo under the:  
*no*  
+ shifts gaze to YU
- 19 (0.5)
- 20 °er° şey °er:° ne [diyelim]  
*like what shall we say*
- 21 YU: [şöyle ] dicem ben  
*this is what I will say*  
+ shifts gaze back to his paper
- 22 =who [helped] (.) jo:↑hn
- 23 ME: [king ]
- 24 (0.3)

25 YU: er::  
+ ME shifts gaze to her paper

26 (0.3)

27 ↑with his soldiers dicem  
*I will say*  
+ turns gaze to ME  
+ ME shifts gaze to YU's paper

28 (0.3)

29 hani jeremy's soldiers vardı ya:  
**well you know there was "jeremy's soldiers"**

30 ME: ↑a: olabilir aslında (.) ne demiştin  
**actually it can be possible, what have you said**  
+ turns gaze to YU, then to her paper  
+ YU shifts gaze back to his paper  
+ positions hand on her paper in an attempt to write the new description

31 =is?

32 YU: is (0.2) [the s↑p:y ]

33 ME: [the soldier ↓m1]  
**is it "the soldier"**

34 (0.3)

35 1# is the sp:y  
+ YU shifts gaze to ME's paper, then back to his paper

36 (0.4)

37 YU: who helped (.) john (.) with his soldiers  
+ turns gaze to ME's paper

38 ((YU shifts gaze back to his paper and then to ME's paper again while ME continues to make the change for 3.0s))

39 ME: °güzel°  
**good**  
+ YU shifts gaze back to his paper

40 (1.5)

41 YU: hani soldiers'ın<sup>#1</sup> nolduğunu (.) hatırlarlar  
**well they would remember what "soldiers" is**  
+ ME crosses out the replaced name  
+ shifts gaze to ME, then back to his paper  
#1(lines 35-41) ME writes down the new sentence as YU articulates it

42 ME: aynen (.) [jeremy ]  
**exactly**  
+ writes down the new character name "Jeremy" in place of "Lizzie"

43 YU: [>huh-huh<]

44 (1.0)

45 ME: tamam  
**okay**

Extract 10b begins with ME's preface marked by a turn-initial pitch rise which projects the upcoming suggestion and decision making process for the problematic character name "Lizzie" and her initiation of a new sequence. ME's preface is met with YU's agreement to initiate a change process for the matching item in the IR interaction through his use of the compliance token "evet/*yes*". Looking at his exam paper, YU then immediately offers a suggestion that he has thought of for a character name to be tested as a replacement of the problematic name "Lizzie" just as ME requests a suggestion from YU in line 4 with a shift of gaze to YU's exam paper. YU offers the name "Jeremy Stickles" in line 5 by underlining this name given in the box on his paper. Following the discourse marker "*hani/well*", an elongated hesitation marker as well as an intra-turn pause, YU then starts to formulate a description for the suggested character name.

YU articulates the first part of the description with a rising intonation and an elongation on the second syllable of the word "*s ↑ p : y*" in line 5. He continues to formulate the rest of the description on spot between lines 6 and 9 as evident in the intra-turn pauses and an elongated hesitation marker that follow. As he is searching for a description, he completes his turn slowly and in line 7, utters the first part of the relative clause that defines the spy that he has already announced (i.e., *who helped*). This is received with ME's elongated change-of-state token "*↑ hnm :*" (Heritage, 1984b) uttered with a pitch rise in word-initial position in line 10 just as she orients to her paper with a pen in her hand in an attempt to write what YU has just produced, which displays her interest in YU's suggestion and acknowledges it. ME's change-of-state token also shows that YU's suggestion creates a change in her epistemic state with regard to its relevance to the matching test as a replacement of the problematic character name "Lizzie".

Following a gap of silence, YU continues with his turn by repeating the previously produced component of the relative clause (i.e., *who helped*) in line 12 accompanied by a thinking face as he is still searching for a way to complete the relative clause. After a gap of silence, he completes the relative clause with the name "*j oh n*" in line 14 with an emphasis on its final sound. At this point, ME starts to write the new description on her paper but then stops to offer a different suggestion for the relative clause in line 16 before YU completes his description of the character name. Producing a thinking face, ME provides the alternative relative clause "*who worked (.) <for the sake of>*" in a slow pace but it is

incomplete as she is still in the process of formulating it. This draws YU's attention as he shifts his gaze to ME. ME's online formulation of her relative clause continues in line 18 following a gap of silence and an elongated hesitation marker where she performs a self-initiated self-repair for the expression "for the sake of" in the relative clause she has suggested and replaces it with the repair solution "under the:" by using the repair preface "yo/no" (Lerner & Kitzinger, 2010). At this point, she shifts her gaze to YU and then launches a word search in line 20 through her stretching of the final word of her turn (i.e., the:), intra-turn pause and her use of hesitation markers as she is having problem in producing a linguistic item to complete her relative clause. Failing to come up with an expression, ME requests YU's help using an inclusive language (ne diyelim/*what shall we say*).

Overlapping ME at a turn-final position, with a shift of gaze back to his paper, YU takes over in line 21 by using a preface that projects continuation of his turn which was already in progress before ME's interference to offer her alternative. His use of "will" as well as the first person singular pronoun in his preface also displays certainty about his suggestion for the new character description (şöyle dicem ben/*this is what I will say*). YU's self-selection technique using a preface also shows that he knows what to say about the suggested character but is in the process of finding a way to express what he has in mind. Immediately following his preface, he, thus, repeats his prior not-yet-completed turn where he framed the first component of the defining relative clause (i.e., who helped (.) jɔ: ↑hn) by producing "jɔ: ↑hn" with a rising intonation and an elongation in line 22 and then after some gaps of silences and his elongated hesitation marker, he eventually completes and finalizes his prior relative clause with "↑with his soldiers". In doing this, he initially raises the pitch of his turn, emphasizes the first two words and once again uses "will" as a marker of certainty (dicem/*I will say*), which all show that he has come up with what he has been looking for. Meanwhile, just as YU takes over to repeat his prior relative clause, ME overlaps and completes her relative clause with "king" in line 23, resolving her word search and then shifts her gaze away from YU to her paper as YU does not orient to her alternative but attempts to complete the relative clause he has started to formulate. As YU articulates the final component of his relative clause in line 27, he shifts his gaze to ME while ME turns her gaze to YU's paper, orienting to his suggestion.

Following a gap of brief silence with no uptake from ME, YU provides account for his suggestion in line 29 by making reference to “*Jeremy’s soldiers*” in the story using the reference marker “*ya:/you know*” with an elongation. In this way, he backs up his suggestion and provides a justification for it with reference to shared knowledge that is assumed to have been gained in a past event, that is the act of reading the story, which is the shared product, before the IR session (Can Daşkın & Hatipoğlu, 2019; You, 2015). In line 30, this is met with an immediate responsive uptake by ME, who displays an upgraded agreement and an affiliative stance with YU through her use of the elongated turn-initial preface “*↑a:*” uttered with a pitch rise and her acceptance token “*olabilir/it can be possible*”. Her turn-initial preface can also be defined as a change-of-state token (Heritage, 1984b) and a response cry (Goffman, 1981) that demonstrates ME’s recall of the shared information from the story book and her transfer of her knowledge about Jeremy into the current context (i.e., the matching test) upon YU’s account. Just as ME responds to YU’s suggestion, she also shifts her gaze to YU, maintaining mutual gaze with him for a short time and then both ME and YU turn their gazes back to their papers.

Positioning her hand with a pencil on her exam paper in an attempt to write the newly offered description, ME requests YU’s repetition and dictation of his character description quickly followed by her DIU (Koshik, 2002) where she reads aloud the first word “*is?*” in the description that is already written on her paper with an interrogative intonation in line 31 to elicit YU’s repetition of his suggestion. As a result, YU restates the first part of the description “*is the s↑p:y*” in line 32 in alignment with ME’s DIU with a pitch rise and an elongation on the second syllable of “*s↑p:y*”. Before YU articulates the rest of the description, ME herself responds to her DIU in line 33 by employing an anticipatory completion (Lerner, 1996a) in the form of a confirmation request, checking her recall of the candidate suggestion in an overlap with YU’s turn. With YU’s repetition of his turn (*is the sp:y*) in line 35 to make it hearable by ME, ME’s anticipatory completion and thus, her recall of YU’s description turns out to be wrong since she refers not to the “*spy*” but the “*soldier*” that appears in the relative clause of the description. Just as YU takes a turn to repeat his turn, ME starts to write down the new item on her exam paper at YU’s dictation, which further displays her uptake of YU’s suggestion and their ultimate decision on the change.

As the collaborative writing activity takes place between lines 35 and 41, YU shifts his gaze to ME's paper and then back to his paper three times, checking to see if the suggested sentence is written correctly. While ME is making the change on her paper after YU's articulation of the relative clause in the item, ME quietly offers a positive assessment of the change in line 39. Following a lapse of 1.5 seconds, YU initiates a non-minimal post-expansion sequence (Stivers, 2013) in line 41 upon the group's agreement and decision on the change. In this post-expansion sequence, he provides further account for the new character description but this time, he frames his account from the students' perspective by orienting to the violated testing principle and assuming the students' recall of the information about the soldiers mentioned in the story book (i.e., *soldiers' in nolduğunu (.) hatırlarlar/they would remember what "soldiers" is*). By once again making reference to shared knowledge about the story, YU shows that the new character "Jeremy" is more likely to be remembered by the students as opposed to the problematic character "Lizzie" and thus, proves that the problem identified in the matching test (i.e., the item tests a minor detail from the story that is unlikely to be remembered by the students) is solved.

Through the end of YU's turn of account, having completed her writing of the new description on her paper, ME crosses out the problematic name "Lizzie" and then writes "Jeremy" in its place accompanied by her articulation of it in line 42 while YU shifts his gaze to ME and then back to his paper. YU's account receives ME's acknowledgement in line 42 as ME uses the compliance token "aynen", which is, in turn, further acknowledged by YU in overlap with ME's turn in line 43. Following a gap of silence, ME then proposes sequence closure via her use of the minimal acceptance token "tamam/okay".

Extract 10b demonstrates the suggestion and decision making process to change the previously problematized character name "Lizzie" and its description in a matching test of a story exam where the students are to match the character names with their descriptions. In order to replace this item with a major character name that is more likely to be remembered by the students, YU offers the name "Jeremy Stickles" and immediately starts to formulate a description for it spontaneously. Since the item is to be changed completely, the group needs to decide on two things: (1) the character name that is to be tested in replacement of the problematic character "Lizzie", and (2) a description for the selected character. The character "Jeremy" suggested by YU is directly accepted by the group as the interaction



mainly focuses on the formulation of a description for this character without further work and discussion on character selection. However, the two participants offer two alternative relative clauses that define “Jeremy”, which focus on the different aspects of this character. The two suggestions for the new character description provided by YU and ME respectively are (1) *who helped John with his soldiers*, and (2) *who worked under the king*. The group orients to the first description provided by YU and finds it more relevant to the matching test, eventually deciding on it. This is evident in ME’s change-of-state tokens (Heritage, 1984b), her display of upgraded agreement, positive assessment of the suggested change, and her noting of the change on her paper.

There are certain interactional actions that prompt the group to decide on YU’s suggestion. First of all, as in the most of the other extracts, YU’s account-giving seems to play an important role in the acceptance of his suggestion. It is only after YU’s justification for the character description he has formulated that ME shows an explicit agreement and does not insist on her alternative or does not continue to provide any other alternatives. YU backs up his suggestion by making reference to the shared product (i.e., the story book), which the participants both have equal epistemic access to and thus, to shared knowledge that is assumed to have been gained in a past event, that is the act of reading the story before the IR session (Can Daşkın & Hatipoğlu, 2019; You, 2015). Such an interactional act of account-giving helps ME recall the shared information from the story book and creates a change in her epistemic state with regard to the relevance of YU’s suggestion to the matching test as a replacement of the problematic character name “Lizzie”. Using a change-of-state token (Heritage, 1984b), she, thus, transfers her knowledge about Jeremy into the current context (i.e., the matching test) and then marks the suggested change on her exam paper, exhibiting acceptance of it and the group’s ultimate decision. As a result, this extract shows that the suggestion that is justified based on evidence is likely to be accepted and put to use by the group.

Although ME’s display of an upgraded agreement and affiliative stance with YU does not project an expansion, YU even initiates a non-minimal post-expansion sequence where he provides further account for his suggestion, making his point clearer. Different from his prior account, this time he frames his account from the students’ perspective by orienting to the violated testing principle and assuming the students’ recall of the information about the soldiers mentioned in the story book. By once again making reference to shared knowledge

about the story, YU shows that the new character “Jeremy” is more likely to be remembered by the students as opposed to the problematic character “Lizzie” and thus, proves that the problem identified in the matching test (i.e., the item tests a minor detail from the story that is unlikely to be remembered by the students) is solved. In this way, he makes sure that the ultimate goal of the change process is achieved and mutual understanding is established in the group.

Secondly, YU’s display of a stance of certainty with regard to his suggestion throughout the talk-in-interaction might have a role in the acceptance of his suggestion. Although YU formulates the new character description on spot along with hesitation markers and intra-turn pauses, it is clear that he knows what to say about the character Jeremy but just searches for a way to express what he has in mind. This is because he had some time to think about the item before the IR interaction while taking the test individually and hence, it is seen that as soon as ME initiates the new sequence, he immediately offers a character name even before ME’s request for a suggestion from him and articulates the first component of his description. As a result, he sticks to his description throughout the interaction and does not even orient to ME’s alternative. This becomes more apparent especially in his preface in line 21 where he projects continuation of his ongoing turn of suggesting a description before ME’s interference to offer her alternative and displays certainty about his suggestion by using “*will*” and the first person singular pronoun (*şöyle dicem ben/this is what I will say*). Similarly, immediately following the formulation of his description in line 27, he uses “*will*” and the first person singular pronoun again as a marker of certainty (*dicem/I will say*) and provides a justification for his suggestion in line 29. This further conveys a stance that embraces his suggestion and encourages its uptake by the group. Furthermore, while articulating his description, YU frequently uses a rising intonation, an emphasis, and an elongation, which are the other interactional resources that display a stance of certainty with respect to YU’s selection of words that express what he has in mind.

Regarding the institutional role distribution in extract 10b, there are indications of tester, reviewer and teacher roles and responsibilities in this IR interaction. To begin with the tester roles, ME orients to her epistemic right, responsibility and obligation (Heritage, 2013) to monitor the change process as the tester by initiating a sequence of suggestion-making for change in the test through her use of a preface and then her request for a suggestion from the

reviewer YU. It is only after ME's preface and her display of a stance of willingness to change the item that YU offers his suggestion. ME also affects the decision-making process by showing an uptake of the offered change through an upgraded agreement and her change-of-state tokens (Heritage, 1984b) and closing the sequence through her noting of the new item on her paper and her positive assessment of the change. Additionally, ME ensures correct understanding of the change by constructing a DIU (Koshik, 2002) to elicit YU's repetition of his suggestion before making the change on her paper and employing an anticipatory completion (Lerner, 1996a) in the form of a confirmation request. In fact, her recall of the description YU has offered turns out to be wrong in her anticipatory completion, which shows that the tester's final check in and through talk with respect to her understanding of what is suggested is important in establishing shared understanding. In this way, ME orients to her role as the tester who is accountable for the test.

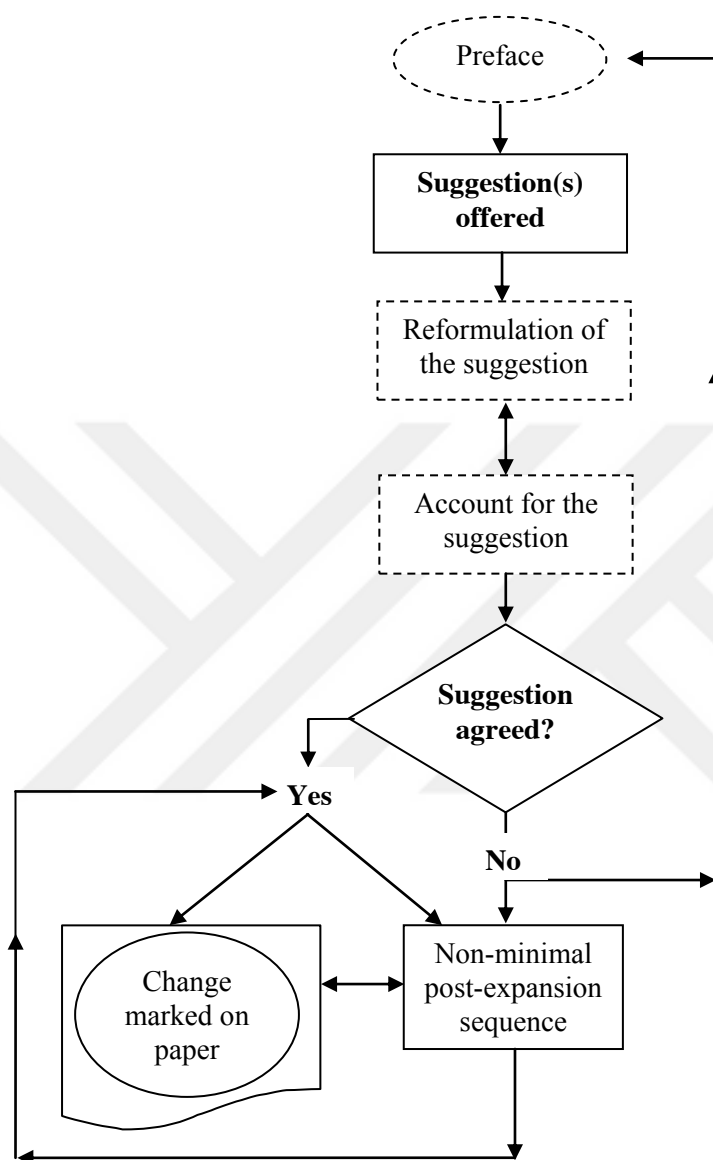
With respect to the reviewer roles, ME ascribes YU an epistemic right and responsibility to come up with an alternative by requesting a suggestion from him regarding the character name to be tested in replacement of "Lizzie". In this way, she orients to the reviewer's role as a "suggestion-maker" during the phase of SC in the IR interaction. For the same reason, YU uses the first person singular pronoun "ben" when providing his suggestion, which shows that as the reviewer, he has done his job of reviewing the item and considering a solution for the identified problem. In addition, while ME is making the change on her paper, YU cross checks the description by shifting his gaze to ME's paper and then back to his paper three times. In this way, he makes sure that the suggested sentence is understood and written correctly, which signals his right, responsibility and obligation (Heritage, 2013) to ensure shared understanding as the reviewer of the test and as the provider of the suggestion. Finally, YU provides a justification for his suggestion by making reference to the story book, which demonstrates that YU has fulfilled his responsibility to read the story before coming to the IR meeting, orienting to his K+ status about the issue (Heritage, 2013). As for the teacher roles in extract 10b, the reviewer YU orients to his institutional role as a teacher by framing his second account from the students' perspectives. This implies his epistemic right, responsibility and obligation (Heritage, 2013) to review the item with consideration of the needs and abilities of the target group of students.

Despite an orientation to the institutional role distribution of the participants, a collaborative process is followed via some interactional resources during the change process in the

problematic item of the matching test. Firstly, the participants both seek a description for the suggested character “Jeremy” and try to formulate an alternative spontaneously, which is evident in their elongations, hesitation markers, intra-turn pauses, gaps of silences, repetition of the previously uttered components of the character description, self-initiated self-repair and a thinking face. For instance, while YU is searching for a way to complete his defining relative clause, ME interferes in an attempt to help YU formulate a description for Jeremy and considers a different alternative that might change the trajectory of the interaction. She does not just wait for YU to make a suggestion but gets engaged in the change process through a collaborative effort. Likewise, following ME’s word search and display of hesitation as she is trying to come up with an alternative relative clause, YU takes the floor in line 21 to change the focus of the interaction to his suggestion, stopping ME to make a further attempt to find an alternative. In this way, the two participants monitor the change process collaboratively until they reach a mutually acceptable final resolution in replacing the problematic item about “Lizzie”.

Secondly, ME invites a collaborative suggestion-making process by requesting YU’s help with her word search using an inclusive language in line 20 (*ne diyelim/what shall we say*). Similarly, ME initiates a collaborative writing activity via her DIU and anticipatory completion in the form of a confirmation request. As a result, she writes the new item on her paper as YU articulates it one by one. This establishes mutual understanding in the group with respect to the decided change in the test. Lastly, the participants’ shared responsibility, and equal and active engagement in the phase of SC is evident in their gaze behaviour as they continuously look at their own or each others’ exam papers throughout the talk-in-interaction (Goodwin, 1981; Rossano, 2013). Such a collaborative process can be explained in terms of the participants’ assumption of equal epistemic state and authority since they all have equal epistemic access to the story book and can offer or evaluate a suggestion based on the information in this story.

#### 4.2.1. Summary of Interactional Patterns in Phase SC



\*See Appendix C for the meaning of flowchart symbols

**Figure 4. Overall Structural Organization of Phase SC**

Regarding the findings for the overall structural organization of the phase of SC in the language test IR interactions analysed for this study, it is seen that this phase is about the participants' suggestion and decision making process for a change to overcome the

identified problem in the item, which is found to be operationalized through a particular recurring interactional pattern as illustrated in Figure 4. First of all, before a suggestion is offered, as seen in all the extracts except for extract 3b, the participants use a preface that projects the upcoming suggestion-making process for the problem in the item. The preface serves several functions in the talk-in-interaction in the phase of SC. It enables the participants to (1) initiate the phase of SC by projecting the change process for the revision of the problematized item as the relevant next action in the IR interaction (extracts 1b, 2b, 5b, 6b, 7b, 8b, 9b, 10b) (e.g., *↑belki bire başka bişey/maybe for one something else*, extract 10b), (2) elicit suggestions for the problematized item as the sequence-initiating action (extracts 1b, 2b, 4b, 7b, 9b, 10b) (e.g., *hani (.) onun yerine ne diyebiliriz acaba/well I wonder what we can say instead of that*, extract 1b), (3) project the announcement of a suggestion that one has in mind or is in the process of formulating and thus, to draw the other participants' attention to this upcoming suggestion (extracts 1b, 4b, 5b, 6b, 8b, 10b) (e.g., *şöyle diyelim/let's say this*, extract 4b), and (4) launch new suggestion-making sequences within the phase of SC, seeking different alternatives for the problematized item (extracts 4b, 7b) (e.g., *°burda daha ne diyebiliriz ki°/what else can we say here*, extract 7b).

The preface is then followed by the participants' offer of a suggestion in the base FPP where one of the participants suggests a change in the item to overcome the problem announced in the first phase of PI. At the core of phase SC is suggesting a change to solve the problem in the language test item, which is a significant interactional action emerging in all the extracts analysed for this study. It is this process of suggestion-making that especially shapes the talk-in-interaction in phase SC and demonstrates how the participants interactionally deal with the problem identified in the first phase of the IR interaction with respect to the language test items that are reviewed. As Table 14 shows, it is found that in instances where multiple suggestions are offered before deciding on a change in the problematized item (extracts 1b, 4b, 7b), one suggestion paves the way for another suggestion in the talk-in-interaction as the participants are constantly in search of a better alternative for change in the item. With the formulation and evaluation of each suggestion, the participants make an attempt to produce a better one that does not have the weaknesses of the previous ones. This shows that the articulation of every new idea that comes to mind and its discussion in and

through talk is important in improving the suggestions and giving rise to better alternatives that eventually lead to a mutually acceptable change.

In relation to this, it is seen that in those instances, what makes the decided alternative more acceptable by the participants is that it is different from the other suggestions and thus, changes the trajectory of the problematized part of the item. To illustrate, in extract 1b, four different suggestions are provided to replace the trouble source “*experts*” in the correct option and the decided alternative is different from the prior suggestions as well as what is stated and problematized in the original option (i.e., *experts*) in that it is not a person designing the special diet mentioned in the option but it is about the content of the diet, making it more thought-provoking and plausible as the correct answer as desired by the group. Likewise, in extract 4b, seven different suggestions are offered to reconstruct option “b” so that it is incorrect as a distractor and the group decides on the final one since it deviates from the original option by presenting a completely different content, changing the trajectory of the problematized option. When the first suggestions fail in eliminating the problem in the item or raise other problems, the group realizes that the problematized option requires a greater change in terms of its content. Such a process demonstrates the role of interaction in testing all the possibilities in the decision-making process as it gives a chance for immediate feedback until a mutually acceptable change that addresses the problem in the item is made.

**Table 14. Summary of Suggestions and Decided Changes for the Problems in the Items**

| Extract No | Problem in the Item                                                                                                                                                                                                                                                                                                    | Suggestions                                                                                                                | Decided Change                                         |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------|
| 1b         | The correct option “b” in a multiple-choice reading comprehension item is too obvious and too easy for the students to select as an advantage <i>not</i> mentioned in the reading text.<br><br><b>Trouble source:</b> “ <i>experts</i> ” in the correct option “ <i>following a special diet designed by experts</i> ” | 1. “ <i>professionals</i> ”<br>2. “ <i>older people</i> ”<br>3. “ <i>the elderly</i> ”<br>4. “ <i>consisting of meat</i> ” | “ <i>following a special diet consisting of meat</i> ” |

**Table 14 (continued)**

|    |                                                                                                                                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                  |                                                                        |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|
| 2b | The verb “ <i>cover</i> ” given in a tense passage to be transformed into its past tense form is not studied in class and the students may not know and understand it.<br><br><b>Trouble source:</b> “ <i>cover</i> ”                                                                        | “ <i>change</i> ”                                                                                                                                                                                                                                                                                                                                                                | “ <i>change</i> ”                                                      |
| 3b | The distractor “ <i>therefore</i> ” in a multiple-choice cloze test is also possible as a correct answer besides the expected answer “ <i>furthermore</i> ”.<br><br><b>Trouble source:</b> “ <i>therefore</i> ”                                                                              | “ <i>whereas</i> ”                                                                                                                                                                                                                                                                                                                                                               | “ <i>whereas</i> ”                                                     |
| 4b | The distractor option “b” in a multiple-choice reading comprehension item is also possible as a correct answer.<br><br><b>Trouble source:</b> option “b”- “ <i>customers do shopping very fast</i> ”                                                                                         | 1. changing the adverb “ <i>fast</i> ” into its opposite “ <i>slowly</i> ”<br>2. making the option negative ( <i>don’t do shopping</i> )<br>3. “ <i>customers wait for shop...</i> ”<br>4. “ <i>customers buy nothing at all</i> ”<br>5. “ <i>people feel better</i> ”<br>6. “ <i>customers prefer natural music</i> ”<br>7. “ <i>customers prefer more expensive products</i> ” | “ <i>customers prefer to buy expensive products</i> ”                  |
| 5b | The time expression “ <i>before</i> ” used in a multiple-choice cloze test item as a prompt for the past perfect tense tested does not fit well in the sentence and is inappropriate within the given context.<br><br><b>Trouble source:</b> “ <i>before</i> ”                               | 1. “ <i>in their early school</i> ”<br>2. “ <i>in their early education time</i> ”<br>3. “ <i>in their early education life</i> ”<br>4. “ <i>early years of their education</i> ”                                                                                                                                                                                                | “ <i>the early years of their education</i> ”                          |
| 6b | The expression “ <i>physical reactions</i> ” is also possible as a correct answer besides the expected answer “ <i>chemicals</i> ” in one of the fill-in-the-blanks given in a lecture outline for a note-taking listening quiz.<br><br><b>Trouble source:</b> “ <i>physical reactions</i> ” | 1. including “ <i>physical reactions</i> ” in the item as a continuation of the available sentence<br>2. giving both alternative answers in the answer key separated by a slash                                                                                                                                                                                                  | giving both alternative answers in the answer key separated by a slash |



**Table 14 (continued)**

|     |                                                                                                                                                                                                                                            |                                                                                                                                                                                                                                           |                                                                         |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| 7b  | The distractor “ <i>might</i> ” in a binary-choice item given in a text is also possible as a correct answer besides the expected answer “ <i>should</i> ”.                                                                                | 1. “ <i>have to</i> ”, “ <i>must</i> ”<br>2. “ <i>shouldn’t</i> ”<br>3. “ <i>don’t have to</i> ”                                                                                                                                          | “ <i>don’t have to</i> ”                                                |
|     | <b>Trouble source:</b> “ <i>might</i> ”                                                                                                                                                                                                    |                                                                                                                                                                                                                                           |                                                                         |
| 8b  | The students might not be able to understand the verb “ <i>underrecognize</i> ” given in the part of the text that a multiple-choice reading comprehension item is based on due to their lack of knowledge of the prefix “ <i>under</i> ”. | 1. “ <i>neglected</i> ”<br>2. “ <i>failed to recognize</i> ”<br><br>For corrections in the sentence:<br>Object of the sentence:<br>1. “ <i>people</i> ”<br>2. “ <i>scientists</i> ”<br>Subject of the sentence:<br>“ <i>mould cases</i> ” | “ <i>scientists have failed to recognize mould cases</i> ”              |
|     | <b>Trouble source:</b><br>“ <i>underrecognize</i> ” in the sentence “ <i>moulds have been an underrecognized health problem</i> ”                                                                                                          |                                                                                                                                                                                                                                           |                                                                         |
| 9b  | The distractor option “ <i>d</i> ” in a multiple-choice reading comprehension item is also possible as a correct answer.                                                                                                                   | 1. making the option negative<br>2. “ <i>should</i> ”                                                                                                                                                                                     | “ <i>should</i> ”                                                       |
|     | <b>Trouble source:</b><br>The modal verb “ <i>can</i> ” in the option “ <i>people can freely criticize Michael when he does not play</i> ”                                                                                                 |                                                                                                                                                                                                                                           |                                                                         |
| 10b | The character “ <i>Lizzie</i> ” tested in a matching test of a story exam is a minor character which the students might have difficulty remembering.                                                                                       | For “ <i>Lizzie</i> ”:<br><i>Jeremy Stickles</i><br><br>For the new character description:<br>1. “ <i>is the spy who helped John with his soldiers</i> ”<br>2. “ <i>is the spy who worked under the king</i> ”                            | “ <i>Jeremy Stickles is the spy who helped John with his soldiers</i> ” |
|     | <b>Trouble source:</b> “ <i>Lizzie</i> ” and its description                                                                                                                                                                               |                                                                                                                                                                                                                                           |                                                                         |

In extracts where two or multiple suggestions are offered (See Table 14), the analysis has further shown that two suggestions might be made subsequently for change in the problematized item (extracts 1b, 4b, 6b, 8b, 9b, 10b). That is, two of the participants provide their suggestions respectively without orienting to the one that has already been made. This could be either because they already have a suggestion in mind to share with the group or

because they are after better alternatives, proceeding with their search for an alternative, which can be interpreted as a display of dispreference with the prior suggestion that has been offered but that has not been oriented to. Only after they reveal their suggestions, they orient to one of them and make an evaluation of it, eventually deciding on a change to revise the item. For instance, in extract 9b, when GE provides the first suggestion of making the option negative, the group does not orient to it but as they keep their gazes at the test material, SA comes up with another alternative (i.e., replacing the trouble source “*can*” in the distractor option with “*should*”), which the group directly focuses and decides on. Likewise, in extract 8b, in order to replace the trouble source “*underrecognize*” that is claimed to be difficult for the students to understand due to their lack of knowledge of the prefix “*under*”, two suggestions are offered respectively for the group to consider (i.e., *neglected*, *failed to recognize*) and the group decides on the second one provided by ME (i.e., *failed to recognize*).

On the other hand, as Table 14 illustrates, there are instances where the IR group focuses on one particular suggestion that is ultimately accepted (extracts 2b, 3b, 5b, 10b). In those instances, either the offered suggestion is directly accepted (e.g., “*whereas*” in extract 3b; “*Jeremy Stickle*” in extract 10b) without seeking other alternatives, or the talk-in-interaction revolves around its evaluation (extracts 2b, 3b) or the reformulation of the suggested expression (extract 5b), as a result of which the group makes a decision in the end accordingly. For instance, in extract 2b, the verb “*change*” is suggested as an alternative verb for the trouble source “*cover*” given in a tense passage to be transformed into its past tense form and the group carefully evaluates this verb in terms of the students’ familiarity with it before displaying an uptake of it. In extract 5b, too, one particular time expression is suggested as a possible replacement of “*before*” given in the target sentence as a prompt for the past perfect tense tested in a multiple-choice cloze test and the group focuses on this expression, together reformulating it so that it fits better in the sentence within the given context in the text and the problem is solved.

Regarding the interactional act of suggesting a change to overcome the problem in the item, the analysis has further revealed that one change made in the item might require another change as in extracts 8b and 10b (See Table 14). In those instances, once the group decides on a major change that addresses the trouble source in the item, they are seen to be in a position to proceed with the change process in the rest of the interaction by initiating a new

suggestion-making sequence for a different part of the item to make the change in the item complete. To illustrate, in extract 8b, the major change decided for the problematized verb “*underrecognize*” given in the target sentence in the reading text (i.e., *failed to recognize*) requires other changes and corrections in the subject and object of the sentence so that the newly suggested verb expression “*failed to recognize*” that fulfils the group’s expectations fits meaningfully and accurately in the sentence. Consequently, another suggestion and decision making process starts for the changes in the subject and object of the sentence. In the end, the original sentence “*moulds have been an underrecognized health problem*” is eventually transformed into “*scientists have failed to recognize mould cases*” where the subject and the object of the sentence as well as the form of the sentence from passive to active are changed and corrected along with the main change in the verb expression. Such a process demonstrates that the participants should be alert to any other problems that might emerge after a particular change is made in the test material.

Likewise, in extract 10b, in order to replace the item with a major character name from the story book that is more likely to be remembered by the students, the group needs to decide on two things respectively: (1) the character name that is to be tested in replacement of the problematized character “Lizzie”, and (2) a description for the selected new character. Hence, the group’s decision to test the character “Jeremy Stickles” in place of “Lizzie” inevitably leads to a new suggestion-making process to formulate a description for this new character so that the revision of the matching test item can be complete. The group needs to make a joint decision on the new character description as well to ensure that it presents a better option than the original item and does not raise problems that could go unnoticed. As a result, the findings show that the change process in the revision of a problematic item in an IR interaction can be much more complex and dynamic, requiring a careful and critical eye.

Considering the suggestions offered for change in the problematized items in the extracts analysed for this study, it is found that they do not come as ready-made solutions but are often formulated spontaneously in and through the talk-in-interaction. This is especially apparent in extracts 1b, 4b, 5b, 7b, 8b, and 10b where the participants make their online thinking process visible while searching for or formulating a suggestion spontaneously by means of interactional resources such as (1) hesitation markers often appearing in an elongated form (extracts 1b, 4b, 7b, 10b), (2) a preface that projects their upcoming suggestion (extracts 1b, 4b, 5b, 10b) (e.g., *şey olur mu?/would this thing be*

*alright?*, extract 1b), (3) gaps of silences, intra-turn cut-offs and pauses (extracts 1b, 4b, 5b, 7b, 10b), (4) repeating the same words or expressions previously uttered (extracts 4b, 10b) (e.g., customers (.) pre↑fer (.) u:hm ↑natural (0.4) and natural music, extract 4b), (5) starting to read aloud the sentence in question and then stopping before the trouble source (extracts 1b, 8b) (e.g., following a special diet-, extract 1b), (6) redesigning one's suggestion (older people of the (0.3) tribe, extract 1b), (7) withdrawing the suggestion (extracts 1b, 4b) (e.g., buy nothing at all *desek bu sefer de direk* (.) çok (.) şey çelişcek/*if we say "buy nothing at all", this time then it will directly like contradict a lot*, extract 4b), (8) doing silent reading (extract 7b), (9) inserting the suggestion into the target sentence and reading it aloud in L1 (i.e., Turkish) (e.g., <dikkatli olmak> ve onları korumak zorunda değiller/*the british don't have to be careful and protect them*, extract 7b), (10) performing a self-initiated self-repair (who worked (.) <for the sake of> er: yo/no under the:, extract 10b), and (11) launching a word search (°er° şey/*like* °er:° ne diyelim/*what shall we say*, extract 10b). Those might be accompanied by the participants' non-verbal actions such as producing a thinking face (Goodwin & Goodwin, 1986) (extracts 1b, 4b, 7b, 10b), continuously looking at their exam papers (extracts 1b, 4b, 7b, 8b, 10b), shifting their gaze between the reading text and the item (extracts 1b, 4b), searching through the textbook (extract 7b), mouthing some words (extract 7b), and scanning through the sentence in question (extract 8b). Such a process of spontaneous formulation of a suggestion demonstrates that the participants are actively engaged in the phase of SC whether their suggestions are then reformulated, accepted, rejected and self-cancelled, shaping the talk-in-interaction considerably. Even though those suggestions are produced spontaneously in and through talk without too much consideration beforehand, this seems to have interactional implications as they enable the participants to approach the item cautiously with a critical eye and give immediate feedback, leading to the emergence of new ideas and discussions in the course of solving the problem in the item.

With respect to the interactional act of suggesting a change for the problematized item, as Table 15 illustrates, it is found that the participants produce their suggestions by either downgrading them through, for example, hedges and confirmation requests (extracts 1b, 2b, 4b, 5b, 6b, 7b, 9b) or displaying a stance of certainty towards them through, for instance,

their positive claims and assertions about their suggestion in declaratively-formatted turns (extracts 3b, 4b, 6b, 8b, 10b). In the case of some suggestions, although the participants are seen to initially downgrade their suggestion, they then turn out to display certainty towards it (extracts 1b, 7b, 9b) or the other participants approach the offered suggestion with certainty (extracts 2b, 5b). The analysis has shown that the participants' downgrading of their suggestions and/or display of a stance of certainty with regard to their suggestions have some interactional implications, playing an important role in the joint decision-making process. It is revealed that the epistemic downgrading of a suggestion paves the way for further collaborative thought on the suggested alternative, leading to either its acceptance (extracts 1b, 2b, 5b, 7b, 9b) or rejection (extracts 4b, 6b, 7b) through the group's negotiation and discussion on it. On the other hand, the display of certainty with regard to a suggestion in the first place makes it more likely to be accepted by the group (extracts 3b, 4b, 6b, 8b, 10b). When the suggestion that is initially presented with an epistemic downgrading is eventually accepted, it is also the suggestion-maker's and/or the other participants' subsequent stance of certainty while discussing and evaluating it that seems to play an important role in the group's uptake of the suggestion (extracts 1b, 2b, 7b, 9b).

**Table 15. Interactional Resources for Downgrading or Displaying Certainty towards a Suggestion**

| No                                               | Downgrading a Suggestion                                           | Extract No                      | Examples                                                                                                                                                                                                 |
|--------------------------------------------------|--------------------------------------------------------------------|---------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1                                                | Confirmation requests                                              | 1b, 4b, 6b, 7b, 9b              | shouldn't çok mu basit<br>o↓lur/would "shouldn't" be too easy (line 14, extract 7b)                                                                                                                      |
| 2                                                | Hedges and expressions of uncertainty                              | 2b, 5b                          | belki change geçmiştir diye düşündüm emin değilim ama/I thought maybe "change" appeared but I'm not sure (lines 3, 6, extract 2b)                                                                        |
| 3                                                | Presentation of the suggestion as an example                       | 4b, 5b                          | gibi bişiy (.) desek/what if we say something like (line 36, extract 4b)                                                                                                                                 |
| <b>Displaying Certainty towards a Suggestion</b> |                                                                    |                                 |                                                                                                                                                                                                          |
| 4                                                | Positive assertions about the suggestion in declarative statements | 1b, 2b, 4b, 6b, 7b, 8b, 9b, 10b | böyle bi bilgi yok/there is no such information (line 50, extract 1b)<br>o zaman should koyduğumuz za↑man (.) yanlış oluyo sık/then when we put "should", the option becomes wrong (line 15, extract 9b) |

**Table 15 (continued)**

|   |                                               |                 |                                                                                                                  |
|---|-----------------------------------------------|-----------------|------------------------------------------------------------------------------------------------------------------|
|   |                                               |                 | öyle bişey yok/ <i>there is no such thing</i> (line 66, extract 4b)                                              |
| 5 | Reassertion of a claim about the suggestion   | 2b, 4b, 6b, 7b  | korumak zorunda değiller niye desin/ <i>why would one say "they don't have to protect"</i> (line 66, extract 7b) |
| 6 | Higher volume and/or emphasis                 | 3b, 4b, 8b, 10b | ↑ <u>WHEREAS</u> (line 1, extract 3b)                                                                            |
| 7 | The stance marker " <i>tabi / of course</i> " | 1b, 2b, 6b      | tabi daha iyi bilirler change'i/ <i>of course they would know "change" better</i> (line 7, extract 2b)           |

To exemplify, in extract 2b, while YE's downgrading of her suggestion through hedges (See Table 15.2) paves the way for further collaborative thought on the suggested verb "*change*", CI's stronger evaluative stance that the students would certainly know the verb "*change*" as well as her reassertions of her account (i.e., *evet change'i daha iyi bilirler/yes they would know "change" better*) and use of the stance marker "*tabi / of course*" (Stivers, 2011) lead to its immediate acceptance by the tester SA. Upon CI's evaluation, the appropriacy of the verb is assured in terms of the testing principle that states "the item should represent a specific content area covered in the course" (i.e., "test what you teach") (Carr, 2011; Haladyna & Rodriguez, 2013) and the verb is not further questioned, being directly substituted with the problematized one on the exam paper without a need to seek other alternative verbs. In extract 7b, too, CI downgrades her suggestions throughout the talk-in-interaction by presenting them in the form of a confirmation request with an orientation to an assumed shared testing principle (See Table 15.1) and hence, invites collaborative inquiry into her suggestions. However, upon the tester AY's hesitation about CI's last suggestion, CI then conveys a stance of certainty towards her suggestion by giving thorough account and concrete evidence from the text with a stronger evaluative stance (e.g., *korumak zorunda değiller (.) yanlış oluyo/"they don't have to protect" is wrong*), which seems to play an important role in the acceptance of her suggestion by AY.

Since the downgrading of a suggestion makes it available for negotiation and opens up a group discussion on it, it might not always be accepted (extracts 4b, 6b, 7b). In extract 4b,

for instance, two of the offered suggestions are downgraded through confirmation requests and their presentation just as one example for the group to consider (i.e., people feel better olur mu?/would it be okay?, suggestion 5; customers (.) pre↑fer (.) u:hm ↑natural (0.4) and natural music gibi bişiy (.) desek/what if we say something like that, suggestion 6), which prompts a response and an evaluation of the presented suggestions from the other participants. However, the group rejects those suggestions by challenging them and making a negative evaluation of them, leading to the rise of new suggestions.

On the other hand, in some of the extracts, it is seen that the suggestions that are decided for change in the problematized items are presented with a stance of certainty in the first place. To illustrate, in extract 3b, the problem is dealt with in a more straightforward way since one particular suggestion is offered by the reviewer DE (i.e., *whereas*) with a stance of certainty and confidence as evident in her articulation of her suggestion at a higher volume and with an emphasis (i.e., ↑WHEREAS) and then the suggestion is directly accepted by the other participants as an alternative for the problematic distractor “*therefore*”. DE’s display of certainty does not explicitly place responsibility on the other participants to evaluate the offered suggestion and thus, without overtly providing account for it, the participants show an acceptance of the suggestion “*whereas*” after quickly evaluating it individually by considering the context given in the cloze text with their gazes already at the item. Considering the decided suggestion in extract 4b, too, SA displays a stance of certainty towards her suggestion by uttering the word “*more*” at a higher volume with an emphasis on it, giving account for it in a declaratively-formatted turn (i.e., öyle bişey yok/*there is no such thing*), and reasserting her claims for the suggestion throughout the talk-in-interaction (e.g., çünkü öyle bişey yok aslında/*because there is actually no such thing*), which might all have an important role in the group’s uptake of and decision on SA’s suggestion.

Before a response uptake is displayed with regard to the offered suggestion(s) in the base SPP, as an optional step appearing in the third position in Figure 4, a suggested expression might undergo a reformulation through a collaborative process by the IR group as in extracts 1b, 4b, and 5b. In those instances, some slight changes and modifications are made in the new alternative by the other participants before they evaluate it in terms of its appropriacy

and role in eliminating the problem in the item and before they display a stance of (dis)agreement with respect to it. As seen in extracts 1b, 4b, and 5b, an alternative expression that is suggested by one participant can be reformulated and redesigned through the talk-in-interaction in an attempt to improve it, address the problem adequately and make it mutually acceptable. It is found that the interactional resources that play a role in the reformulation of a suggested expression are the participants' repair actions, DIU's, spontaneous formulation of a suggestion, anticipatory completions, collaborative turn sequences, and embedded correction, which shows that the suggestions are collaboratively shaped in and through talk.

To illustrate, in extract 1b, the suggestion "*the elderly*" emerges as the reformulation of the previous suggestion "*older people*" through EK's repair action and the use of a DIU (i.e., *designed by?*). Likewise, EK's repair initiation (i.e., *°nasıl?°/what?*) enables DE to redesign her prior suggestion "*older people*" as "*older people of the tribe*" in an attempt to make it more meaningful and understandable by the other participants. Similarly, in extract 4b, the final suggestion that is accepted, in fact, emerges as the reformulation of the prior one since SA picks the first component of the previously suggested sentence (i.e., *customers prefer*) and reformulates it by completing it with a different piece of information. In this respect, the previously offered suggestion that was rejected leads to a new idea in and through talk, illustrating the importance of interaction during the suggestion and decision making process. The accepted suggestion then undergoes a second collaborative reformulation through the other participants' repair initiation (e.g., *nasıl?/sorry?*) and GE's use of a DIU (i.e., *customers-*).

Finally, in extract 5b, the talk-in-interaction centres around the reformulation of one particular time expression suggested by MI as a rephrase of the trouble source "*before*" in the target sentence. Every time the suggested expression is articulated, it becomes subject to small changes, eventually being finalized and decided on. The reformulation process starts with MI's initiation of her spontaneous formulation of the expression as her word search mobilizes ME's anticipatory completion, which then leads to the first reformulation of the expression as MI's "*in their early school*" becomes "*in their early education time*" through collaborative turn sequences. Then ME's failure to recall the expression announced by MI prompts MI's repair action and embedded correction, which in turn leads to the second reformulation of the expression as "*in their early education time*" is transformed into "*in*



*their early education life*". This is then finalized by AV's restructuring and rephrasing of the expression with "*early years of their education*" and its insertion into the target sentence by ME as "*in the early years of their education*". As a result, "*in their early school*" initially articulated just in the process of being formulated is eventually completed and changed into "*in the early years of their education*" through a collaborative process and met with the group's uptake.

As another optional step appearing in the fourth position in Figure 4, the base FPP of offering a suggestion might also be followed by the suggestion-maker's account-giving for his/her alternative preceding the base SPP, which appears in all the extracts except for extracts 1b and 3b. As illustrated by the double-sided arrow in Figure 4, if applicable, the reformulation of a suggested expression might also take place after the act of account-giving (extract 4b) or continue following it (extract 5b). Thus, the position of those two optional steps (i.e., reformulation and account-giving) can change. With respect to the interactional act of account-giving in phase SC, the suggestion-makers justify their suggestions based on evidence before relinquishing the floor and making relevant next the other participants' display of stance towards the suggestion. The findings show that the suggestion that is justified based on evidence is more likely to be accepted and put to use by the group.

As presented in Table 16, the participants often make a thorough evaluation of an offered suggestion and provide account for it from different perspectives and dimensions, which affects the decision-making process to a considerable extent. They, for instance, present evidence from the available shared sources such as the reading text, listening notes, and the story book as external sources of authority (Jacknick, 2013) (e.g., *aslında babası >insanlar criticize< etsin demiyo kötü bişey diyo/because his father doesn't actually say people shall criticize, he says it is something bad*, extract 9b) and make reference to past instructional events based on a shared curriculum (Can Daşkın & Hatipoğlu, 2019; You, 2015) (e.g., *belki change geçmiştir diye düşündüm/I thought maybe "change" appeared*, extract 2b), the students' present state of L2 knowledge and skills (e.g., *daha iyi bilirler change'i/they would know "change" better*, extract 2b) as well as their level of proficiency (e.g., *hani belki biraz da zorlamak/well maybe to also make it a bit difficult*, extract 5b) in order to prove the

appropriacy and validity of the suggestion in terms of eliminating the problem identified in the language test item. In their accounts, the participants are also found to translate a suggested expression or the target sentence in an item into their L1 (i.e., Turkish) with the suggested change inserted in it so as to clarify the meaning of the newly formulated expression and make a connection between the information in the text and the suggested change (extracts 5b, 7b).

**Table 16. Summary of Accounts and Evidence Presented for the Suggestions**

| Extract No | Suggestions                                                                                                                           | Account/Evidence                                                                                                                                                                                                                                         |
|------------|---------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2b         | <i>change</i>                                                                                                                         | The verb “ <i>change</i> ” appeared in the course and the students know it better.<br>(Test what you teach principle)                                                                                                                                    |
| 4b         | <i>customers prefer more expensive products</i>                                                                                       | The reading text includes no such information that would make the distractor option correct. In the text, it is explained that muzak soundscapes make people buy products but not necessarily the expensive ones.<br>(Only one correct answer principle) |
| 5b         | <i>the early years of their education</i>                                                                                             | Such an expression in place of “ <i>before</i> ” in the target sentence would fit better in the sentence and make the item more difficult.<br>(Appropriacy of language principle, Level of difficulty principle)                                         |
| 6b         | including “ <i>physical reactions</i> ” in the item as a continuation of the available sentence                                       | This would prevent the students from writing “ <i>physical reactions</i> ” as an alternative answer in the blank.<br>(Sufficiency of context principle, Only one correct answer principle)                                                               |
|            | giving both alternative answers (i.e., “ <i>chemicals</i> ” and “ <i>physical reactions</i> ”) in the answer key separated by a slash | Both alternatives are possible as correct answers.<br>(Identification of acceptable answers principle)                                                                                                                                                   |

**Table 16 (continued)**

|     |                                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-----|-----------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 7b  | <i>don't have to</i>                                                                                      | The reason-result relation in the surrounding context is not meaningful when the distractor " <i>don't have to</i> " is inserted into the positive verb expressions (i.e., <i>be careful, protect</i> ) in the target sentence in the binary-choice cloze test item but at the same time such a distractor is more plausible within the context in the text, making the item more thought-provoking for the target group of students. Plus, the students should know the verb " <i>protect</i> ".<br><br>(Only one correct answer principle,<br>Plausibility principle,<br>Simplicity of vocabulary principle) |
| 8b  | <i>scientists</i> (for the correction in the subject of the sentence that is revised in the reading text) | This would correspond with the expression " <i>health care professionals</i> " used in the text.<br><br>(Text coherence principle)                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 9b  | people <i>should</i> freely criticize Michael when he does not play                                       | As explained in the reading text, the father's view of the public criticisms directed at his son as one of the disadvantages of being famous makes the distractor option where " <i>can</i> " is replaced with " <i>should</i> " unlikely as a correct answer. Thus, the father's suggesting the public to criticize his son as claimed in the new distractor option is not the point in question in the text.<br><br>(Only one correct answer principle)                                                                                                                                                      |
| 10b | <i>Jeremy Stickles is the spy who helped John with his soldiers</i>                                       | The students would be able to remember what " <i>soldiers</i> " is.<br><br>(Avoidance of trivial content principle)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

As explained in Table 16, while providing an account for a suggestion, the participants happen to address the violated testing principles that were revealed in the phase of PI (See Section 4.1.1), further proving their suggestion as a valid solution for the identified problem in the item. To exemplify, in extracts 4b, 7b, and 9b, the suggestions offered for the possible correct distractors are justified in terms of the violated testing principle that each item should have only one correct answer (Heaton, 1990; Hatipoğlu, 2009) (e.g., *should* koyduğumuz za↑man (.) yanlış oluyo sık/*when we put "should", the option becomes wrong*, extract 9b). Similarly, in extract 2b, the reviewers address the violated testing principle that the item should represent a specific content area covered in the course (i.e., test what you teach) (Carr, 2011; Haladyna & Rodriguez, 2013)

by explaining that the suggested verb “*change*” appeared in the course and the students are more familiar with it than the original verb “*cover*” (e.g., *change’i daha iyi bilirler/they would know “change” better*). In extract 5b, an orientation is shown to the violated testing principle that the language used in the item should sound as natural as possible (Harris, 1969) by AV’s candidate explanation for the replacement of “*before*” in the target sentence (i.e., *güzel otursun cümle içine diye mi/is it because so that it fits better in the sentence*). Finally, in extract 10b, while providing account for the new character description, the reviewer YU makes reference to the violated testing principle that each item should be based on important content (Haladyna & Rodriguez, 2013) by claiming that the students would remember this information from the story (i.e., *soldiers’ın nolduğunu (.) hatırlarlar/they would remember what “soldiers” is*) unlike the problematized character description which tested a minor detail that was difficult to recall from the story.

The analysis with regard to the interactional act of account-giving has also shown that the participants orient to not only the violated testing principles previously identified in phase PI but also other testing principles that the offered suggestions would simultaneously fulfil, displaying their situated understanding of how a test should be (See Table 16). For instance, in extract 7b, while the main issue is the possibility of the distractor as a correct answer, CI considers her suggestion “*don’t have to*” not only in terms of its unlikelihood as a correct answer within the context given in the item and hence, the testing principle that each item should have only one correct answer but also from the students’ perspectives. In doing so, CI explains that the suggested distractor modal verb (i.e., *don’t have to*) would make the students think more about the item without directly selecting the correct answer “*should*” (i.e., *don’t have to biraz daha düşündürebilir onları/“don’t have to” could make them think a little bit more*). This displays an orientation to CI’s situated understanding of how a test should be and thus, to an assumed shared testing principle concerning the plausibility of the options in selected response items (i.e., options should be reasonably attractive and plausible). She, thus, goes beyond the violated testing principle and makes sure that her suggestion fulfils other testing principles regarding the quality of options in a multiple-choice item.

In the same extract, by explaining that the students should also know the verb “*protect*” given in the target sentence, which she regards as one piece of evidence for the unlikelihood of the new distractor “*don’t have to*” as a correct answer (i.e., *protect’in (.) de anlamını <bilmeleri> gerekir/they should also know the meaning of “protect”*), CI orients to the testing principle that the vocabulary in a text should be kept simple and appropriate for the test-takers so that “reading comprehension does not interfere with testing the content intended” (Haladyna, 2004, p.99). Therefore, CI at the same time makes sure that the target sentence poses no problems for the test-takers and is clear enough for them to be able to show their knowledge of modal verbs studied in class; thereby, establishing the purity of the item.

Similarly, in extract 5b, it is not only the announced testing principle concerning the appropriacy of the language used in the item but also the testing principle about the difficulty level of the item that is addressed in the accounts given for the new expression. MI explains that the replacement of “*before*” with the suggested expression “*the early years of their education*” would make the item relatively more difficult for the intended group of test-takers in terms of its linguistic complexity (i.e., *hani belki biraz da zorlamak/well maybe to also make it a bit difficult*). Thus, she makes reference to the testing principle that the item should be at a level appropriate to the proficiency level of the test-takers (Heaton, 1990). In this way, she also displays her situated understanding of how a test item should be through her accounts for the suggestion.

Additionally, in extract 6b, while providing account for the suggestions offered for the other possible correct answer (i.e., *physical reactions*) in a fill-in-the-blank item in the lecture outline, the participants refer to some other testing principles described in the literature for fill-in items (Brown, 1996). SA, for instance, suggests including “*physical reactions*” in the item as a continuation of the available sentence by explaining that it would prevent the students from writing it as an alternative answer in the blank. In this way, she displays an orientation to the testing principle that there should be “sufficient context to convey the intent of the question to the students” in a fill-in item (Brown, 1996, p.58). This would, in turn, ensure that the required response is concise (Brown, 1996) and there is only one correct answer (Heaton, 1990; Hatipoğlu, 2009). However, SA at the same time considers whether the inclusion of extra information in the item would extend the sentence (i.e., *>size*

ama< uzatmış mı oluruz cümleyi/*but do you think we would end up having extended the sentence*) since “extra context will burden students with extraneous material to read and may inadvertently provide students with extraneous clues” (Brown, 1996, p.59). This addresses the testing principles that excessive verbiage should be avoided in the item and the stem should not contain extraneous information or irrelevant clues (Haladyna, 2004; Hatipoğlu, 2009; Heaton, 1990).

In the same extract, DU makes another suggestion, that is giving both alternative answers (i.e., “*chemicals*” and “*physical reactions*”) in the answer key separated by a slash. This not only acknowledges that there is more than one possible correct answer (e.g., *ikisi de doğru/both are correct*) but also regards the announcement of the problem as an opportunity rather than a violation of a testing principle (i.e., each item should have only one correct answer) to provide a list of acceptable answers for the blank. As an acknowledged approach for productive response items, this would ensure that the answer key is complete (Brown, 1996). Overall, it is seen that through their accounts, the participants show that their suggestion addresses not only the problem identified in the item but also other assumed testing principles that would further improve the validity and reliability of the item.

In one instance as in extract 6b, the analysis has further revealed that the participants’ interactional act of account-giving for a change from different perspectives plays an important role in prompting a repair action and clearing up a misunderstanding about the suggested change (i.e., *↑huh şey ↑key’e ikisini de vereyim/well I shall give both of them in the key*). In this respect, it is seen that account-giving serves several functions, which can be summarized as follows: (1) justifying a suggestion based on evidence, (2) addressing testing principles for item writing, (3) ensuring that the problem is adequately addressed, (4) displaying a stance of certainty towards an offered suggestion (See Table 15.4 and 15.5), (5) establishing shared understanding in the group with respect to the suggested change, and (6) clearing up a misunderstanding about the suggested change.

Once the suggestion is made followed by its reformulation by the group and/or the suggestion-maker’s account-giving where necessary, the interactional process is unfolded in two ways; first when the group agrees on the offered suggestion, launching the base SPP

(labelled as “yes” in Figure 4) and second when the suggestion-makers continue with their turn and challenge their own suggestion themselves, which then leads to the self-cancellation and withdrawal of the suggestion, or when one of the other participants displays disagreement in the base SPP by challenging and questioning the appropriacy of the suggestion (labelled as “no” in Figure 4).

In the first case, the group directly agrees with the first offered suggestion or with one of the two suggestions offered subsequently (extracts 2b, 3b, 5b, 6b, 8b, 9b, 10b). Once the participants agree and decide on a suggestion, they then make the necessary changes on their exam papers, which further shows their uptake of the suggested change, and the item is eventually revised in order to make it more valid and free of problems for the target group of students. At this stage, while making the change, a collaborative writing activity might take place as in extracts 1b, 8b, and 10b where the participants note down the decided change on their exam papers by inserting the accepted change into the item/text and reading aloud the revised part of the item/text as a whole through collaborative turn completions. As in extracts 4b, 5b, 6b, 7b, and 8b, the tester herself might also read aloud the modified part of the item/text or restate the kind of change that is decided while or just before noting it down on her paper. All these ensure shared understanding in the group with respect to the change made in the item. In addition, it is found that although the participants’ agreement with the suggestion does not project an expansion in phase SC, it might be followed by a non-minimal post-expansion sequence (Stivers, 2013) after, before or simultaneously with the noting of the decided change on the exam paper where the participants further elaborate on the suggested change, making sure that the change addresses the announced problem and establishing mutual understanding in the group with respect to the change made in the item (extracts 2b, 4b, 6b, 9b, 10b).

In the second case, a suggestion is challenged and a stance of disagreement is displayed with regard to it (extracts 1b, 4b, 7b). In challenging an offered suggestion, the participants present evidence from the available shared sources (e.g., reading text) and make claims with reference to the violated testing principles announced in the phase of PI as well as other testing principles that the suggestion would further violate as presented in Table 17. They are also found to challenge not only the other participants’ suggestions but also their own suggestions, which results in the withdrawal and self-cancellation of one’s own suggestion. For instance, in extract 1b, DE approaches the suggestions for the problematized correct

option cautiously and challenges her suggestion “*older people*” (i.e., AMA Bİ DAKKA olderlar (0.3) yapar mı acaba/*but wait a minute, I wonder if the old would do it*) as well as EK’s suggestion “*consisting of meat*” with respect to their availability in the text so that the option does not end up being expressed as an advantage in the text and thus, being an incorrect option since the item asks for an advantage *not* mentioned in the text. Thus, while challenging those suggestions, DE addresses the testing principle regarding items with “no answer” as a result of revisions and rewrites in the item (Hatipoğlu, 2009) rather than the plausibility of the suggested options as the main concern, which seems to be already ensured with the selection of new content that is more relevant to the topic of tribes explained in the text (i.e., *older people, meat*). In this extract, DE also challenges her first suggestion “*professionals*” (i.e., olmaz (.) gene/*still not*) upon her realization that it is similar to the original option, retaining the violated testing principle that the option is too obvious for the students to select. In this way, as in the case of the problematized option, she makes an orientation to the violation of the testing principle that “options should be reasonably attractive and plausible” (Carr, 2011; Haladyna & Rodriguez, 2013; Hatipoğlu, 2009; Heaton, 1990), showing that her suggestion does not address the identified problem in the item.

Similarly, in extract 4b, five of the seven suggestions for the problematized distractor option are challenged and decided not to be acted on. The tester SA and the reviewer DU are active in doing this as they challenge either their own suggestions or the others’ suggestions by providing account and evidence for their unacceptability as a distractor with a consideration of the context given in the reading text and an orientation to the testing principles for multiple-choice item construction. To exemplify, the participants challenge the suggestions “*customers do shopping very slow*” and “*people feel better*” by claiming that they present information that is already available in the reading text (e.g., >dördüncü paragraftan ötürü< doğru olur/*because of the fourth paragraph, it would be correct*; o da ama doğru işte bi şey geneline baktığınız zaman parçanın/*but it is also something right you see when you look at the overall text*), which once again makes the distractor option correct, failing to overcome the identified problem and fulfil the violated testing principle that each item should have only one correct answer (Carr, 2011; Haladyna & Rodriguez, 2013; Hatipoğlu, 2009; Heaton, 1990).



As for the suggestions “*buy nothing at all*” and “*customers prefer natural music*” in extract 4b, the participants challenge them in terms of their plausibility rather than the violated testing principle that the item should have only one correct answer. SA claims that “*buy nothing at all*” would directly contradict the information in the reading text, turning out to be too obvious to eliminate and thus, ending up as an implausible distractor (e.g., *direk (.) çok şey çelişcek =çok belli olacak/it will directly like contradict a lot, it will be too obvious*). Likewise, she explains that “*customers prefer natural music*” is not meaningful with the stem of the item that states “*when Muzak soundscapes are used*” (e.g., *ama muzak (.) soundscape olduğun↑da (.) onlar müzik tercihi yapmış olmuyo ki:↓/but when it is “muzak soundscape”, they do not happen to have made a music preference*). This, too, addresses the violation of the testing principle that options should be reasonably attractive and plausible so that the “students obtain the correct option by direct selection rather than the elimination of obviously incorrect options” (Heaton, 1990, p.32). In addition, the suggestion “*don’t do shopping*” is challenged in terms of the “odd man out” principle (Carr, 2011; Hatipoğlu, 2009) as it would turn out to be the only negative option in the item (i.e., *tek er olumsuz şık o olur/it would be the only negative option*), violating the testing principle that “options should be similar in structure” (Fulcher, 2010, p.172) and homogenous in grammatical structure (Haladyna & Rodriguez, 2013). This is because such an option would end up being immediately noticed as the “odd man” and discarded by test-takers (Hatipoğlu, 2009).

Finally, in extract 7b, the act of challenging a suggestion takes place three times throughout the talk-in-interaction until a decision is made on the change for the distractor option in a binary-choice item. While the first considered alternative modal verbs “*have to/must*” are challenged by the suggestion-maker CI herself in terms of their unacceptability as a distractor due to their close association with the expected correct answer “*should*” (i.e., *have to must falan hep riskli oluyor zaten/“have to” “must” and so on happen to be always risky*), the following two suggestions “*shouldn’t*” and “*don’t have to*” are challenged by the tester AY, who points out that “*shouldn’t*” makes no sense within the given context (i.e., *dikkatli olmamalısın pek olmuyo sanki/“you shouldn’t be careful” seems to be unlikely*), and questions “*don’t have to*” in terms of its possibility as a correct answer

within the context given in the cloze test (i.e., uyuyo mu? peki/*but then does it fit in*). Consequently, while the suggestions “*have to/must*” and “*don’t have to*” are challenged with respect to the violated testing principle that each item should have only one correct answer and hence, whether they provide a solution for the identified problem, the suggestion “*shouldn’t*” is questioned in terms of its plausibility as a distractor rather than its correctness since the target sentence becomes too meaningless when “*shouldn’t*” is inserted into it and read along with the verb expression “*be careful*”, which is a positive action. Although the incorrectness of the distractor “*shouldn’t*” is ensured as intended, it is still not accepted due to its implausibility, which is another testing principle to be considered when writing selective response items. As a result, it is seen that as in the case of account-giving, the participants approach the suggestions critically and challenge them in terms of their failure not only in eliminating the identified problem in the item but also in fulfilling other testing principles for writing selective response items.

**Table 17. Summary of Challenges Presented to the Suggestions and Their Outcomes**

| Extract No | Suggestions Challenged        | Presented Challenge                                                                                                                                          | Outcome                                                      |
|------------|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|
| 1b         | “ <i>professionals</i> ”      | Making the option still too obvious to identify for the students’ level<br>(Plausibility principle)                                                          | Suggestion self-cancelled                                    |
|            | “ <i>older people</i> ”       | The possibility of being available in the text as the question asks for information <i>not</i> mentioned in the text<br>(Avoidance of “no answer” principle) | Challenge rejected / Suggestion considered as an alternative |
|            | “ <i>consisting of meat</i> ” | The possibility of being available in the text as the question asks for information <i>not</i> mentioned in the text<br>(Avoidance of “no answer” principle) | Challenge rejected / Suggestion accepted                     |

**Table 17 (continued)**

|    |                                                                         |                                                                                                                                                        |                                          |
|----|-------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|
| 4b | Changing the adverb “ <i>fast</i> ” into its opposite “ <i>slowly</i> ” | Being available in the reading text, making the distractor option possible as a correct answer<br><br>(Only one correct answer principle)              | Suggestion self-cancelled                |
|    | Making the option negative ( <i>don’t do shopping</i> )                 | Making the option the only negative one<br><br>(The odd man out principle)                                                                             | Challenge accepted / Suggestion rejected |
|    | “ <i>customers buy nothing at all</i> ”                                 | Contradicting the information in the reading text, making the distractor option too obvious to eliminate<br><br>(Plausibility principle)               | Suggestion self-cancelled                |
|    | “ <i>people feel better</i> ”                                           | Being available in the reading text, making the distractor option possible as a correct answer<br><br>(Only one correct answer principle)              | Suggestion self-cancelled                |
|    | “ <i>customers prefer natural music</i> ”                               | Not completing the stem meaningfully<br><br>(Plausibility principle)                                                                                   | Challenge accepted / Suggestion rejected |
| 7b | “ <i>have to</i> ”, “ <i>must</i> ”                                     | Being unacceptable as a distractor due to their close association with the correct answer “ <i>should</i> ”<br><br>(Only one correct answer principle) | Suggestion self-cancelled                |
|    | “ <i>shouldn’t</i> ”                                                    | Making no sense within the given context<br><br>(Plausibility principle)                                                                               | Challenge accepted / Suggestion rejected |
|    | “ <i>don’t have to</i> ”                                                | The possibility of being correct as a distractor option within the given context<br><br>(Only one correct answer principle)                            | Challenge rejected / Suggestion accepted |

As Table 17 further shows, as a result of the act of challenging a suggestion, the participants (1) self-cancel and eliminate their own suggestion by backing down and withdrawing it in the form of a self-initiated self-repair, which often takes place when the participants

formulate a suggestion spontaneously as soon as the idea comes to their mind (extracts 1b, 4b, 7b) (i.e., labelled as “no” in Figure 4), (2) reject an offered suggestion and then initiate a new suggestion-making sequence, proceeding with their search for an alternative and orienting to different suggestions for change in the problematized item (extracts 1b, 4b, 7b) (i.e., arrow moving from “no” to the beginning of phase SC in Figure 4), or (3) accept the offered suggestion by counter-challenging the disagreeing participant in a non-minimal post-expansion sequence (Stivers, 2013) and conjointly providing further account and elaboration in support of the suggestion, which disproves the presented challenge and builds a consensus in the group about the appropriacy of the suggestion in terms of overcoming the problem in the test item (extracts 1b, 7b) (i.e., arrow moving from NMP to “yes” in Figure 4). Thus, in the third option, projecting an expansion, the act of challenging a suggestion enables the participants to further evaluate the change from different dimensions and makes counter-challenge relevant next in response to the disagreeing participant, as a result of which they eventually reach a consensus on the recommended change.

Additionally, as in extract 4b, it is seen that when a suggestion is challenged, the participants might directly start the whole process of suggestion-making from the beginning and orient to a new search for an alternative for the trouble source in the item without an explicit display of agreement with the presented challenge. With regard to the group’s agreement with the challenge presented to a suggestion, it should also be noted that when a suggestion-maker challenges his/her own alternative right after offering it and engages in a self-cancellation of the suggestion by backing down and withdrawing it, it is found to be received with the other participants’ display of agreement rather than disagreement. Their agreement with the presented challenge does not appear as an expansion but as a response to the suggestion-maker in the second position. However, as in some of the suggestion-making cases, only when another participant challenges an offered suggestion in the base SPP, the group’s agreement with the presented challenge is displayed in the third position as a minimal post-expansion (extracts 4b, 7b) followed by a new suggestion-making sequence. On the other hand, their disagreement with the presented challenge is conveyed in a non-minimal post-expansion sequence so as to counter-challenge the claim made against the suggestion and prove the appropriacy of the suggested change (extracts 1b, 7b).

The findings show that the challenge presented to a suggestion is often accepted and thus, the suggestion is directly rejected by the group or withdrawn by the suggestion-maker

without being further argued for since there is clear evidence for its unacceptability and the issue does not require further negotiation. However, it is seen that only in the cases of “*consisting of meat*” (extract 1b) and “*don’t have to*” (extract 7b), the challenge is rejected and the suggestion is eventually accepted since the act of challenging prompts a group discussion on the validity of the suggestion. This is because the issue is about the availability of the newly suggested information in the reading text and requires a careful analysis of the context given in the text, which might be open to different interpretations. In the case of “*older people*” (extract 1b), too, the challenge is rejected and the suggestion is accepted but in this instance, the suggestion-maker herself withdraws her own challenge (i.e., ↓ *yok/no*), leaving her suggestion as one alternative for the group to consider.

The findings reveal that throughout the phase of SC, the participants often go beyond the use of an agreement token and display mutual agreement with and understanding of an offered suggestion or a claim made about it, whether positive or negative, through a collaborative process. As Table 18 illustrates, they achieve this by (1) employing an anticipatory completion (Lerner, 1996a) whereby the participants complete a participant’s ongoing turn of account-giving for a suggestion through collaborative turn sequences by anticipating what is to come next, (2) providing additional account or evidence for a suggested change, (3) offering a candidate understanding (Antaki, 2012; Heritage, 1984b) of a suggestion or a claim made about it, (4) engaging in a collaborative writing activity where the participants note down the decided change on their exam papers by inserting it into the target sentence and reading aloud the modified part of the item/text as a whole through collaborative turn completions, which at the same time enables a final check for the suggestion and ensures shared understanding regarding the change made in the item, (5) using a full or a partial modified repeat (Kim, 2002; Stivers, 2005) of a claim made about a suggestion, (6) offering a (re)formulation of a prior remark made about a suggestion (Heritage & Watson, 1979), (7) providing a candidate explanation for an offered suggestion, (8) engaging in a choral co-production (Lerner, 2002), and (9) the tester reflecting on the item writing process to support a claim made about a suggestion. Through such collaborative turn sequences, the participants display and demonstrate their shared understanding of the suggestions and claims made about them and at the same time, discuss the suggested changes jointly, which affects the decision-making process to a considerable extent in the phase of SC.

**Table 18. Interactional Resources for Displaying Mutual Agreement through a Collaborative Process in Phase SC**

| No | Display of Mutual Agreement                 | Extract No                  | Examples                                                                                                                                                 |
|----|---------------------------------------------|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Anticipatory completion                     | 1b, 4b, 5b, 6b, 7b, 9b, 10b | alma demiyo/ <i>it does not say "don't buy"</i> (lines 75, 77, extract 4b)                                                                               |
| 2  | Additional account/evidence                 | 2b, 4b, 6b, 9b              | çünkü <kimyasal> şeyler fiziksel reaksiyon oluyo zaten/ <i>because chemical things already happen to be physical reaction</i> (lines 45, 48, extract 6b) |
| 3  | Candidate understanding                     | 2b, 6b                      | verb ikisini bilemezler (.) mi acaba diye diyosunuz/ <i>you say so because they may not know its verb two</i> (lines 11-12, extract 2b)                  |
| 4  | Collaborative writing activity              | 1b, 8b, 10b                 | lines 57-65 (extract 1b)                                                                                                                                 |
| 5  | Full or partial modified repeat of a claim  | 4b, 7b, 9b                  | °tercih yapmıyo°/ <i>they do not make a preference</i> (line 44, extract 4b)                                                                             |
| 6  | (Re)formulation                             | 7b, 9b                      | demiyo/ <i>doesn't say</i> (line 11, extract 9b)<br>°çok saçma di mi°/ <i>it is too nonsense, isn't it</i> (line 21, extract 7b)                         |
| 7  | Candidate explanation for the suggestion    | 5b                          | güzel otursun cümle içine diye mi/ <i>is it because so that you know it fits better in the sentence</i> (line 17)                                        |
| 8  | Choral co-production                        | 1b                          | THE ELDERLY (lines 24-25)                                                                                                                                |
| 9  | Tester's reflection on item writing process | 7b                          | ilk başta must gibi düşünmüştüm/ <i>at first I thought like "must"</i> (line 9)                                                                          |

The findings show that the talk-in-interaction in phase SC displays an orientation to the *institutional role distribution of the participants*. As revealed in this study, the interactional actions emerging in the phase of SC show that the testers and reviewers have certain roles and responsibilities in this phase. To begin with the interactional indicators for tester roles, as illustrated in Table 19, it is seen that as the writers of the test items who are accountable for the test, the testers often check and ensure correct and mutual understanding of the suggested change that is accepted by the group. In doing this, they, for instance, offer a candidate understanding (Antaki, 2012; Heritage, 1984b) of the suggestion in question or

elicit its repetition and dictation through a DIU. In fact, as extract 10b illustrates, such interactional resources used by the tester might help avoid an incorrect noting of the change on the test material. For instance, the tester ME's elicitation of the suggestion through her DIU shows that her recall of YU's alternative for the character description was wrong since she refers not to the "spy" but the "soldier" that appears in the relative clause of the newly formulated description. This indicates that the testers' final check in and through talk with respect to their understanding of what is suggested is important in establishing shared and correct understanding of the suggested change. This is especially essential for accountability issues in testing because in any case of criticism, as the testers of the institution, they will have to explain to the stakeholders why the item is designed in such a way.

Secondly, the testers check shared and correct understanding of the accepted change by reading aloud the modified part of the item/text (e.g., "*prefer to buy expensive products*" in extract 4b) or reannouncing the kind of change that is decided (key'e ikisini de veriyim/*I shall give both of them in the key*, extract 6b) while or just before noting it down on their exam papers. At this point, as in extracts 5b and 8b, they might also insert the suggested change into the target sentence in the item/text and recite it along with the preceding or following words already available in the sentence (e.g., "*in the early years of their education*" in extract 5b, "*have been failed to-*" and "*scientists have failed to recognize mould cases but the situation is changing*" in extract 8b). The testers' reading aloud further functions as a final check for the decided change in the item and hence, solicits mutual understanding in the group with respect to the change made in the item, orienting to their role as the tester who is accountable for the test.

In fact, as illustrated in extract 8b, the testers' act of reading aloud the decided change might lead to the realization of a problem in the revised part of the item and thus, the initiation of a correction sequence. In extract 8b, for instance, just before she makes an attempt to mark the change on her paper, the tester ME's reading aloud of her suggestion "*failed to recognize*" by embedding it in the present perfect passive form given in the target sentence (i.e., have been failed (.) °to°-) enables the reviewer AD to identify the mismatch between the suggested verb expression "*failed to recognize*" and the subject of the sentence "*moulds*". Likewise, ME's reading aloud of the newly formulated sentence through collaborative turn completions with AD (i.e., *scientists have failed to recognize health*

problems) further allows AD to find out that the object of the sentence is also inappropriate after the changes made in the verb and the subject of the sentence. This leads to the initiation of another correction sequence and eventually the careful revision of the target sentence to solve the identified problem.

Thirdly, as illustrated in Table 19.3, in one instance (i.e., extract 7b), the tester AY makes reference to the past item writing process by indicating that “*must*” was an alternative that she thought of while constructing the item but she could not include it as a distractor in the binary-choice item due its close association with the correct answer “*should*”. This signals her right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item as the writer and thus, the ‘owner’ of the item. Finally, the testers’ non-verbal behaviour of noting down the decided change on the test material as seen in all the extracts analysed for this study orients to their responsibility and obligation to finalize the test and make it ready for administration.

**Table 19. Interactional Indicators for Tester and Reviewer Roles and Responsibilities in Phase SC**

| No | Tester Roles and Responsibilities                                                  | Extract No         | Examples                                                                                                                                                       |
|----|------------------------------------------------------------------------------------|--------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Checking understanding of the accepted change through;<br>candidate understanding, | 2b, 6b             | verb ikisini bilemezler (.)mi<br>acaba diye diyosunuz/you say so<br>because they may not know its<br>verb two (lines 11-12, extract 2b)                        |
|    | reading aloud or reannouncement of the change,                                     | 4b, 5b, 6b, 7b, 8b | in the early years of their education (line 35, extract 5b)<br>key’e ikisini de veriyim/I shall give both of them in the key (line 50, extract 6b)             |
|    | elicitation of the suggestion via DIU                                              | 10b                | ne demiştin/what have you said =is? (lines 30-31)                                                                                                              |
| 2  | Noting down the change on the test material                                        | All extracts       | lines 21-26 (extract 8b)                                                                                                                                       |
| 3  | Referring to the item writing process                                              | 7b                 | ilk başta must gibi düşünmüştüm<br>=ama must should’a yakın<br>zaten/at first I thought like<br>“must” but “must” is already<br>close to “should” (lines 9-10) |



**Table 19 (continued)**

| <b>Reviewer Roles and Responsibilities</b> |                                                              |                                                                                                                                                                           |
|--------------------------------------------|--------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Before the IR Session</b>               |                                                              |                                                                                                                                                                           |
| 4                                          | Taking the test as a student                                 | 6b                                                                                                                                                                        |
|                                            |                                                              | bak ben onu yakalamışım/ <i>look I have caught it</i> (line 32)                                                                                                           |
| 5                                          | Reading the story for the reviewing of the story exam        | 10b                                                                                                                                                                       |
|                                            |                                                              | hani jeremy's soldiers vardı ya:/ <i>well you know there was "jeremy's soldiers"</i> (line 29)                                                                            |
| <b>During the IR Session</b>               |                                                              |                                                                                                                                                                           |
| 6                                          | Evaluating an offered suggestion from students' perspectives | 2b, 7b, 10b                                                                                                                                                               |
|                                            |                                                              | daha iyi bilirler change'i/ <i>they would know "change" better</i> (line 7, extract 2b)                                                                                   |
|                                            |                                                              | soldiers'ın nolduğunu (.) hatırlarlar/ <i>they would remember what "soldiers" is</i> (line 41, extract 10b)                                                               |
| 7                                          | Providing further clarification of the problem               | 1b, 5b                                                                                                                                                                    |
|                                            |                                                              | yani ordaki <u>ex</u> ↑ <u>perts</u> (.) ba↑na böyle (.) er: <u>çok</u> -şey geldi/ <i>well to me "experts" there is like too- sounds thing</i> (lines 29-34, extract 1b) |

Regarding the reviewer roles in the phase of SC, as Table 19 demonstrates, the reviewers orient to some of the responsibilities that they should carry out both before and during the IR session as in the phase of PI. To begin with the former (i.e., *before* the IR session), the reviewer DU's reference to her test-taking process by using the past tense or present perfect tense in extract 6b (See Table 19.4) shows that the reviewers are responsible for taking the test as a student would before giving feedback to the tester in the IR interaction. Additionally, in extract 10b, the reviewer YU's justification for his suggestion with reference to the story book by using the reference marker "ya: /*you know*" (See Table 19.5) makes it clear that they come to the IR session for the story exam having already read the story carefully. Those interactional resources indicate that the reviewers have a specific role and responsibility in the IR interactions.

As for the responsibilities to be carried out *during* the IR session, the reviewers mainly show an orientation to their roles as teachers by evaluating an offered suggestion from the students' perspectives as in extracts 2b, 7b, and 10b. This implies the reviewers' epistemic right, responsibility and obligation (Heritage, 2013) to review the item with consideration of

the context of learning as well as the needs and abilities of the target group of students and thus, to create the interplay between teaching and testing. To illustrate, in extract 2b, the evaluation of the suggestion is considered as part of the reviewer responsibility since the state of affair, that is the verbs the students would presumably know, falls into the teachers' epistemic domain as they are best familiar with their students' knowledge of the language they are learning and as part of their institutional roles, they have the right, responsibility and obligation to know and have access to their students' epistemic domain of the target language (Heritage, 2013). This means that unlike the tester SA, the reviewers YE and CI have the right to evaluate the suggestion in terms of their students' knowledge of L2 verbs based on what is taught to them in class so that the item can be co-constructed in line with the testing principle "test what you teach". Similarly, in extract 7b, the reviewer CI orients to her institutional role as a teacher by framing her accounts for her suggestion from the students' perspectives. When providing a justification for her suggestion "*don't have to*", she, for instance, refers to the students with the deictic "*onlari*" and states that the alternative distractor "*don't have to*" could make them think more about the item without directly selecting the expected correct answer "*should*" (i.e., biraz daha düşündürebilir onlari/*it could make them think a little bit more*).

Secondly, the reviewers display an orientation to their roles as problematizers of the item by providing further clarification of the problem in the item during the suggestion and decision making process in the phase of SC (extracts 1b, 5b). In extract 1b, the reviewer BA, who problematized the language test item in the phase of PI, makes this clear by specifically bringing to the fore the trouble source "*experts*" in the correct option and clarifying why she has considered it as problematic (See Table 19.7) while the other participants have already begun to offer suggestions for change. This reminds the group of the matter at hand and ensures shared understanding with respect to why they are to change the option. In this way, as the reviewer of the item, she makes sure that the participants know what to correct in the item and design their suggestions accordingly.

Despite an orientation to the institutional role distribution of the participants, they happen to share the responsibility for the change process in the problematized language test items and for their final decision in such task-oriented IR interactions where the product is shared, affecting everybody in the group. Whether it is the reviewer or the tester, the participants all get actively engaged in the task of suggesting and deciding on a change to overcome the

problem in the item and thus, they all have a contribution to make to this change process. This is evident in the participants' (1) role as suggestion-makers throughout the phase of SC, their constant search and spontaneous formulation of an alternative until they reach a mutually acceptable final resolution, (2) evaluation of a suggestion by, for instance, inserting the suggested change in place of the trouble source in the item and reading it aloud to check its appropriacy before making a comment on it (e.g., inserting the alternative "*should*" into the problematized option and reading it aloud with verb "*criticize*" in extract 9b), (3) display of a stance of (dis)agreement towards an offered suggestion, (4) presentation of a challenge towards a suggested change, (5) account-giving for a suggestion, (6) orientation to testing principles for item construction, (7) cancellation of their own suggestion, (8) projection of the upcoming suggestion-making process through a preface, and (9) alertness to any problems that might emerge in the IR interaction as well as in the change made in the item through, for example, the implementation of a repair action as in extract 6b and the initiation of a correction sequence as in extract 8b. These interactional actions indicate that the participants in the IR interaction are all actively processing as they are suggesting and deciding on a change to improve the item. The involvement of all the participants in suggesting and evaluating a change is especially the case when the identified problem concerns an issue that equally falls within their epistemic domain (e.g., reading text). As a result, the participants assume equal epistemic state and authority since they all put themselves in K+ position as reviewers, teachers and testers who have equal epistemic access to the shared resources and who share a common ground in terms of the institutional expectations with regard to language test construction (Heritage, 2013).

The participants' shared roles and responsibilities can also be explained by the collaborative process that they follow throughout the talk-in-interaction in the phase of SC. They, for instance, (1) invite all the participants' search for an alternative and thus, a collaborative suggestion-making process by applying a preface with an inclusive language, (2) carry out a collaborative inquiry into the offered suggestions, which is most apparent in extract 7b, (3) provide account for a suggestion from different perspectives, which acts as supportive and complementary to the suggestion-maker's explanations, (4) jointly reformulate a suggested expression in collaborative turn sequences, which is most apparent in extract 5b, (5) display mutual agreement with a suggestion or a claim made about it through interactional resources such as anticipatory completion, candidate understanding, and collaborative writing activity (See Table 18), (6) follow a joint decision-making process by inviting collaborative inquiry

into an offered suggestion, eliciting a response from the other participants about a suggested change (e.g., downgrading a suggestion), leaving their suggestions open to negotiation, seeking each others' confirmation, turning the floor over to the other participants, allocating some time for the other participants to think about an offered suggestion as evident in the gaps of silences that follow, and each participant expressing a stance towards a suggestion, and (7) ensure shared understanding with respect to the change by, for example, soliciting and checking the group's confirmation for a suggestion, marking the change on the exam paper through collaborative turn completions, and reading aloud the modified part of the item while noting it down on the test material. All these interactional actions indicate that the change is considered as the product of the whole group.

The participants' equal active engagement in the phase of SC is further evident in their gaze behaviour. It is seen that throughout the interactions, the participants continuously look at their exam papers only sometimes directing their gazes at each other for a short time but this does not indicate disengagement; on the contrary, it signals that the participants actively think about an alternative for the problematized item in such a competing activity (Goodwin, 1981; Rossano, 2013). In addition, the participants' non-verbal actions such as moving the pencil along the problematized part of the item, crossing out the part of the option to be replaced, turning over the page to look at the reading text or the item, producing thinking face, searching through the textbook, mouthing some words, doing silent reading, and marking the change on their exam papers also demonstrate their active engagement in the phase of SC as well as their shared responsibility for the decision and ownership of the test as the shared product, the outcome of which is important for all of them. Finally, the participants' use of an inclusive language throughout the phase of SC in all the extracts once again ascribes an equal epistemic state to the whole group and makes a reference to their shared responsibility for the change process in the item.

### 4.3. Phase 3. Reviewing the Change (RC)

Extract 1c presents the phase of RC in the reviewing interaction on the item problematized in extract 1a (i.e., phase PI) and revised in extract 1b (i.e., phase SC) to overcome the problem. It is to be remembered that in this extract, the two teachers BA and DE, who teach at B2 level, review a multiple-choice reading comprehension item in a level achievement test of B2 groups and provide their feedback to the tester EK. The reading comprehension item has a negative root (i.e., *Which of the following is NOT one of the advantages of the Kalenjini?*) and has four options. In the first segment of the extract (i.e., extract 1a, see Section 4.1), BA problematized the correct option “b” as she believed that it was too obvious and easy to identify for their group of students because the idea of “*experts*” mentioned in the option made it too irrelevant to the content of the reading text, which was about tribes, and could easily be regarded *not* as an advantage of the tribe Kalenjini. Addressing this problem, the option in question underwent change in the second segment of the extract (i.e., extract 1b, see Section 4.2) as the IR group redesigned it through a process of suggestion-making for change and improvement. As a result, the original implausible option “*following a special diet designed by experts*” was replaced with “*following a special diet consisting of meat*” in the phase of SC in extract 1b. However, the reviewing interaction for this particular item does not come to end when a decision is made on the change and the problematic option is revised but continues for further review of the accepted change this time in the following segment of the extract (i.e., extract 1c) to make sure that the change is complete and the problem is solved and that it raises no other problems that could go unnoticed in the item. This extract was also presented at a data session. The analysis provided here is, thus, supported by the analysis carried out at this data session.

#### Extract 1c. Special Diet, Segment c

- 1 EK: ↑böyle mi da[ha iyi?] designed <by'lı mı> kalsın.  
**is it better like this? should it stay with "designed by"**  
+ shifts gaze to DE, then to BA  
+ DE and BA look at their exam papers
- 2 BA: [evet ]  
**yes**
- 3 (0.7)

- 4 °er designed by° =bence ↑böyle daha iyi (.)  
**I think it is better like this**  
+ EK turns gaze to her paper  
+ reads from the option on her exam paper  
+ shifts gaze to EK
- 5 [meat'li daha iyi]  
**it is better with "meat"**  
+ turns gaze to DE
- 6 DE: [consisting of ] meat (.) <sup>1#</sup> evet  
**yes**  
+ shifts gaze to EK  
+ turns gaze back to her paper
- 7 (0.7)
- 8 BA: °meat'li daha iyi°  
**it is better with "meat"**  
+ turns gaze back to her paper
- 9 (0.3)
- 10 DE: evet  
**yes**
- 11 (1.2)
- 12 EK: <sup>#1</sup> meat (.) [dedik di] mi? =milk değil?  
**we said "meat" right? not "milk"**  
+ shifts gaze to DE  
#1(lines 6-12) EK notes the first component of the decided alternative (i.e., *consisting of*) on her paper
- 13 BA: [huh-huh ]
- 14 (1.3)
- 15 [yok (0.2) meat  
**no**  
+ shifts gaze to EK
- 16 DE: [ °meat°  
+ shifts gaze to EK
- 17 EK: [(inaudible) [( (laughs) )  
+ EK writes "meat" on her paper to complete the rest of the new option
- 18 DE: [meat  
+ turns gaze to the reading text on her paper
- 19 (0.2)
- 20 BA: ((laughs))
- 21 EK: yanlış [ °duydum°]  
**I heard it wrong**
- 22 DE: [yoktu ] di mi? [ °hiç o°]  
**it was not present at all was it?**
- 23 BA: [DAIRY ] products (.)  
+ EK shifts gaze to BA

24 er meat and dairy products >diye düşündüm ama  
*I thought so but*  
+ EK turns gaze back to her paper by nodding head  
+ DE shifts gaze to BA

25 [dairy products<  
+ shifts gaze to DE  
+ EK shifts gaze to BA

26 DE: [dairy'i bilemez[↑se  
*what if s/he does not know "dairy"*

27 BA: [er: aslında bilmesi  
28 [lazım tabi çok (.) öğrettik ama]  
*actually s/he should know of course we have taught it a lot but*

29 DE: [bilmesi lazım da: ]  
*s/he should know it but*

30 BA: o ↑da (0.2) sağlıklı bi yiyeceğe: (.) götürebilir (.)  
+ shifts gaze to EK  
+ EK turns gaze back to her paper

31 fikren  
*it then might take us to the idea of healthy food*  
+ shifts gaze to DE

32 DE: huh

33 BA: sağlıklısız bi[şey olsun.  
*it shall be something unhealthy*

34 DE: [er:  
+ produces thinking face with hand on her mouth

35 (0.6)

36 ama meat de o zaman-  
*but then "meat" too-*  
+ EK shifts gaze to DE and then back to her paper  
+ turns gaze back to the reading text  
+ BA turns gaze to her paper

37 ((DE scans through the relevant part of the reading text on her paper with a pencil in her hand for 1.5s while BA and EK look at their exam papers))

38 BA: ↑yok (.) yani o daha böyle şey  
*no well that is more like*  
+ shifts gaze to DE  
+ EK shifts gaze to BA

39 (0.9)

40 °her tarafa çekilebilir bişey. °  
*something that can be interpreted in all ways*  
+ EK turns gaze to her paper

41 (0.4)

- 42 DE: zaten bahsetmi↑yo [ki  
*it does not mention it anyway*  
+ BA turns gaze to her paper
- 43 BA: [tabi tabi bah[setmiyo.  
*of course it does not mention it*  
+ shifts gaze to DE and then back to her paper
- 44 EK: [uh-uh

As the reviewers DE and BA are looking at their exam papers, extract 1c begins with the tester EK's request for final clarification of the change from the other participants through an alternative question by shifting her gaze to DE and then quickly to BA. In doing this, EK refers first to the accepted and noted change (i.e., *consisting of meat*) with the deictic "↑böyle/*like this*", which is marked with a word-initial pitch rise, and then to the prior suggestion that was offered in the phase of SC (i.e., *designed by the elderly*) by reciting its first component as the recognitional reference (i.e., *designed by*), asking and checking the reviewers' preference between these two alternatives. This enables the reviewers to reconsider the alternatives as BA silently reads aloud "°designed by°" from the original option on her exam paper (i.e., *designed by experts*) after a gap of brief silence and her use of a hesitation marker, at which point EK also turns her gaze to her exam paper.

Then shifting her gaze to EK and prefacing her turn by the assessment hedge "bence/*I think*", BA quickly states her opinion and clarifies that the alternative with "*meat*" is a better option by referring to it with the deictic "↑böyle/*like this*" marked with a pitch rise and using the comparative adjective "daha iyi/*better*" in alignment with EK's question. In this way, BA expresses her preference for this already accepted change. In line 5, BA repeats her positive assessment by this time explicitly articulating "*meat*" that is key to the decided alternative with a shift of gaze to DE in order to make her preference with regard to the change in the item clearer. In overlap with BA's turn, shifting her gaze to EK, DE also conveys the same stance in line 6 and displays preference for the alternative with "*meat*" by reading it aloud in full form (i.e., *consisting of meat*) and then using the compliance token "evet/*yes*", which confirms BA. As BA and DE both turn their gazes back to their papers, EK starts to note the change on her paper, which indicates that the alternative with "*meat*" is now accepted as the final resolution and BA repeats her positive



assessment of this alternative with a soft tone of voice in line 8, which is met with DE's confirmation once again in line 10 through her use of the compliance token "evet/yes".

Following a lapse of 1.2 seconds, just as EK completes writing down the first component of the decided alternative (i.e., *consisting of*), she stops before noting the key word "*meat*" on her paper and ensures correct understanding of the alternative by seeking the reviewers' confirmation that it is "meat" and not "milk" in the form of a candidate understanding (Antaki, 2012; Heritage, 1984b) of the change in line 12. She produces "meat" with an emphasis, uses an inclusive language (i.e., *dedik/we said*) and shifts her gaze to DE through the end of her turn-at-talk. This functions as another final check for the decided change before proceeding with the other items on the test material and establishes shared understanding with regard to it. Before EK completes her turn, BA already provides a confirming response for the use of "*meat*" in the new option in line 13 through her acknowledgement token "huh-huh" in overlap with EK's question just after she utters the word "meat".

Following a gap of silence, the reviewers together orient to EK's confirmation-seeking question and confirm her understanding of the change as correct by repeating "*meat*" respectively in alignment through collaborative turn construction with a shift of gaze to EK. Turning her gaze to the reading text on her paper, DE repeats the word the second time in line 18, which further highlights that it is "*meat*" rather than "*milk*". EK responds to the reviewers' confirmation, which aligns with the designed preference of her question, with laughter in line 17. Upon the reviewers' assumed response, EK displays a self-deprecating stance towards her own question through laughter, which signals that her presentation of the alternative "*milk*" was irrelevant suddenly emerging as a new alternative following the decision-making process. As Adelsward and Oberg (1998) point out, laughter is a resource used "in the marking of borders, in the establishing of what is coherent and relevant and what is inconsistent and incongruous" (p.427). Accompanied by her laughter, EK writes "*meat*" on her paper to complete the rest of the new option. BA further displays affiliation with EK by laughing too in line 20. Then EK regards her presentation of another alternative (i.e., *milk*) along with "*meat*" as a hearing problem in line 21 (*yanlış °duydum° / I heard it wrong*), providing an account for her question. This turns out to be the reason why she attempts to check her understanding of the accepted change.

Upon EK's draw of attention to the key word "*meat*" in the new correct option, the review of the change made in the item continues. This time DE plays the devil's advocate in line 22 and checks the other participants' final confirmation with regard to the availability of the information about "*meat*" in the reading text since the item asks for an advantage *not* mentioned in the text. However, the other participants do not orient to her confirmation check; instead, EK's "*milk*" brings a new idea to BA's mind as she then considers "DAIRY products" as an alternative for the problematic option in lines 23 and 24 by producing "DAIRY" at a higher volume and offering it along with "*meat*". This draws the other participants' attention as DE and EK shift their gazes to BA and EK then turns her gaze back to her paper by nodding her head. Yet, BA quickly challenges and withdraws this alternative by launching a *but*-prefaced turn in a fast pace and explaining in lines 30 and 31 that it would be associated with the idea of healthy food, being highly related to the information in the text and thus, ending up as an incorrect answer.

Maintaining mutual gaze with BA, just before BA starts challenging "*dairy products*", DE interferes and overlapping BA's turn-at-talk, she pre-emptively presents another challenge towards this new alternative by questioning "*dairy*" in terms of the students' knowledge of the word with a rising intonation in line 26. BA then cuts off her current unit of talk in line 27 and orients to DE to counter-challenge her by claiming that the word "*dairy*" should be within the students' epistemic domain as it has been taught a lot to them and displaying certainty with her claim through her use of the marked confirmation "*tabi/of course*" (Stivers, 2011) in line 28. BA's evidence prompts DE to quickly back off and change her position in line 29 in an overlapping turn by using a partial modified repeat (Stivers, 2005) of BA's assertion about the students' knowledge of the word "*dairy*" (i.e., *bilmesilazım/s/he should know it*), which not only displays agreement with BA but also makes a counter-claim to her own argument. On the other hand, DE's use of an elongated "*da:/but*" expressing contrast in turn-final position indicates that DE still does not fully agree with BA.

Retaining her floor, BA then carries out her aborted negative evaluation of the new alternative by challenging it in terms of the information in the text rather than the students' knowledge of the word "*dairy*" with a shift of gaze first to EK and then to DE in lines 30 and 31. She then suggests in line 33 that the option should be about something unhealthy so

that it cannot be associated with the advantages mentioned in the text and can be regarded as unstated information, which is what the item asks for. During BA's turn-at-talk, EK keeps her gaze at BA and then turns her gaze back to her paper just prior to BA's completion of her evaluation of the new alternative. In line 32, DE shows agreement with BA via her acknowledgement token "huh".

Upon BA's comment that the special diet mentioned in the option should include an unhealthy food, DE this time challenges the accepted alternative with "*meat*" in terms of its relevance to healthy food in line 36 following her elongated hesitation marker and a brief gap of silence accompanied by her thinking face with her hand on her mouth. Before completing her turn, she cuts off her utterance but the other participants show an orientation to her remark by turning to their exam papers and rechecking the change. While EK and BA look at their exam papers, DE scans through the relevant part of the reading text on her paper with a pencil in her hand for 1.5 seconds. Then in line 38, BA takes over to counter-challenge DE through a turn-initial negative response token marked with a rising intonation (i.e., ↑*yok/no*) and the claim that "*meat*" is something that can be interpreted in all ways, highlighting that it is not directly related to the idea of healthy food. During her turn-in-progress, she shifts her gaze to DE while EK turns her gaze to BA and then back to her paper.

Following a gap of brief silence, DE displays agreement with BA's stance in line 41 by bringing in a supportive and complementary account. She explains that the idea of "*meat*" is not mentioned at all in the reading text. At this point, BA turns her gaze to her paper and then in line 43, shifting her gaze back to DE again, she exhibits agreement and affiliation with DE through her use of the marked confirmation "*tabi/of course*" (Stivers, 2011) articulated twice successively and her partial modified repeat (Kim, 2002; Stivers, 2005) of DE's utterance (i.e., *bahsetmiyo/it does not mention it*). EK further acknowledges DE's idea that "*meat*" is not mentioned at all in the text via her negative response token "uh-uh" in line 44 in an overlapping turn with BA. Eventually, the change decided in the phase of SC is reviewed and confirmed and hence, the language test item is finalized in the IR interaction. As a result, mutual understanding and agreement is established in the IR group with respect to the change made in the item.

Extract 1c demonstrates how the participants review the change that is already decided in the phase of SC to make the correct option plausible and how they finalize the reviewing of the item. This phase can be characterized by the participants' request for final clarification and confirmation checks in an attempt to make sure that the change fulfils the testing principles, causing no additional problems and to establish shared understanding with regard to it before finalizing the item and ending the reviewing interaction for this particular item. The tester EK and the reviewer DE are active in monitoring the review process for the change made in the item. The phase of RC is initiated by EK's request for final clarification of the change through an alternative question (i.e., *↑böyle mi daha iyi? designed <by'lı mı> kalsın/is it better like this? should it stay with "designed by"*) as she checks the reviewers' preference between the accepted change (i.e., *consisting of meat*) and the prior suggestion that was offered (i.e., *designed by the elderly*) but that was not directly rejected, being considered as a possible alternative until a more preferable alternative was suggested. In this way, she draws the reviewers' attention to the prior suggestion that was first met with the group's uptake but then overlooked once they concentrated on the new alternative "*consisting of meat*" and allows them to compare the two possible alternatives so that their decision does not prevent them from missing out a better alternative in terms of its plausibility as requested. As a result, the tester ensures that "*consisting of meat*" is regarded as a better alternative and is thus, more preferable for the reviewers.

Secondly, EK launches a confirmation check in line 12 (i.e., *meat (.) dedik di mi? =milk değil?/we said "meat" right? not "milk"*) in the form of a candidate understanding (Antaki, 2012; Heritage, 1984b) of the change in order to avoid mistakes resulting from possible hearing problems. Through her confirmation check, EK ensures correct and shared understanding of the change so that it can be noted correctly on the exam paper. Lastly, playing the devil's advocate, the reviewer DE also does a final check through her confirmation request in line 22 (i.e., *yoktu di mi? °hiç o°/it was not present at all was it?*). Different from EK's questions, DE does not check the participants' preference for the change or her correct understanding of the word "*meat*" but instead the unavailability of the new information about "*meat*" in the reading text since the item asks for an advantage *not* mentioned in the text. In this way, she makes sure that the new option is not so close to the content of the reading text while trying to make the correct

option more plausible as this might lead to an item without a correct answer, causing more serious problems. Hence, as in the phase of SC, DE once again wants to ensure that the change does not violate other testing principles in an attempt to address the problem identified in the item.

The participants' confirmation checks lead to a display of mutual agreement which is important in establishing shared understanding and making the group's stance clearer so that the change can be overtly confirmed. The participants, for instance, do not just use an agreement token but use such interactional resources as repetitions with full-form recognitional references and modified confirming repeats of a prior utterance as well as marked confirmations. To illustrate, although BA initially responds to EK's request for clarification in alignment with the design of her question by using the deictic "*↑böyle/like this*", she then makes her preference more explicit by using a full-form recognitional reference (i.e., *meat*) in lines 5 and 8 to refer to the accepted alternative and repeating her positive assessment twice (*meat'li daha iyi/it is better with "meat"*), which reconfirms the change in the item. Similarly, DE does not simply use the compliance token "*evet*" but makes her stance clearer by repeating the preferred alternative in full form (i.e., *consisting of meat*) in disalignment with EK's question, going beyond the terms and constraints of her question. Through her non-conforming response (Raymond, 2003) and full-form recognitional reference, she ensures shared understanding with respect to the change that is accepted. Likewise, in response to EK's confirmation-seeking question in line 12, BA and DE provide repetitional responses (Heritage & Raymond, 2012) by explicitly articulating the word "*meat*" to confirm EK that it is "*meat*" and not "*milk*". Although BA initially responds to the question with the negative response token "*yok/no*", she then utters the word "*meat*" to clarify which word she refers to.

Furthermore, DE and BA's partial modified repeats (Kim, 2002; Stivers, 2005) of a prior utterance in lines 29 and 43 respectively (i.e., *bilmesi lazım da:/s/he should know it but; bahsetmiyo/it does not mention it*) display and demonstrate their shared understanding of a claim made in the IR interaction and overtly confirm it. Based on this, by means of such interactional resources, the participants do not only show agreement but also claim their epistemic right to confirm an assertion, indexing epistemic

authority and control over the matter at hand since they all have equal epistemic access to the test material (Heritage & Raymond, 2005; Stivers, 2005). In this way, they “assert the primacy of the second speaker’s rights to make the statement” (Stivers, 2005, p.131). Finally, BA’s use of the marked confirmation “*tabi/of course*” (Stivers, 2011) twice consecutively in line 43 not only exhibits her agreement with DE’s account for the alternative but also displays a stance of certainty towards it, which leads to the acknowledgement of the change in the item and ends the reviewing process.

The analysis has also shown that the phase of RC can include new suggestion-making sequences as in the phase of SC although they do not necessarily result in different changes in the item. Thus, suggestion-making is another interactional action observed in the phase of RC. In this particular extract, EK’s question where she requests confirmation that the main word in the decided alternative is “*meat*” rather than “*milk*” leads to the emergence of a new idea as BA comes up with the alternative “*meat and dairy products*” in and through the talk-in-interaction. This shows that the participants are constantly in search of a better alternative even if they have decided on a change in the problematic item, making the reviewing of the change a dynamic process. However, BA and DE make a quick evaluation of this alternative and find it problematic. Thus, BA backs down and withdraws it, as a result of which the group once again orients to the already accepted alternative “*consisting of meat*”.

As in the first two phases, the act of challenging also appears in the phase of RC in extract 1c. The participants challenge not only the new alternative that arises in RC in and through the talk-in-interaction but also the already accepted alternative as they review them by making claims with reference to the testing principles for item writing. DE employs the first act of challenging in line 26 where she questions the word “*dairy*” in the new alternative offered by BA in terms of the students’ knowledge (i.e., *dairy’ i bilemez/se/what if s/he does not know “dairy”*). In this way, she orients to the testing principle that states “the item should represent a specific content area covered in the course” (i.e., “test what you teach”) (Carr, 2011; Haladyna & Rodriguez, 2013). BA then challenges the idea of dairy products in lines 30-31 as she thinks that it would be associated with the idea of healthy food, being closely related to the information in the text and thus, ending up as an incorrect answer. Although it seems to be a good idea at first sight in terms of the plausibility principle, a closer evaluation of the alternative shows that it would be

problematic, leading to an item with “no answer” as a result of revisions in the item (Hatipoğlu, 2009). BA’s challenging of her alternative from the perspective of healthy food causes DE to approach the accepted alternative “*consisting of meat*” critically and to challenge it this time in terms of its relevance to healthy food mentioned in the reading text (i.e., *ama meat de o zaman-/but then “meat” too-*), which once again addresses the avoidance of “no answer” principle while trying to make the option plausible. Hence, it is seen that one challenge that is presented paves the way for other challenges and opens up new discussions, making the reviewing process of the change effective.

It is found that each act of challenging has different outcomes. BA’s challenging of her own alternative leads to the self-cancellation and withdrawal of “*dairy products*” without any objection from the other participants. On the other hand, DE’s challenges are rejected by BA through her counter-challenges that the word “*dairy*” should be within the students’ epistemic domain and the idea of meat in the new option can be interpreted in all ways, not being directly related to the idea of healthy food. As a result, DE and BA’s challenges entail a further elaboration of the accepted alternative “*consisting of meat*”, enabling the group to achieve consensus on the change and to eventually confirm it “in open discussion, through a process of dialogue and disagreement” (Fulcher & Davidson, 2007, p.21), which makes an important contribution to the validation process.

Regarding the institutional role distribution of the participants in extract 1c, there are indications of tester and teacher roles and responsibilities in this IR interaction. To begin with the tester roles, it is seen that as the writer of the test item who is accountable for the test, the tester EK first checks the reviewers’ preference for the decided change through her request for clarification and then ensures correct and mutual understanding of the accepted change through her candidate understanding (Antaki, 2012; Heritage, 1984b) in the form of a confirmation check. In this way, she reviews the change carefully and receives final confirmation for it before marking it on her exam paper. Additionally, EK’s non-verbal behaviour of noting down the change on the test material orients to her responsibility and obligation to finalize the test item and make it ready for administration.

As for the teacher roles, the reviewers display an orientation to their roles as teachers by evaluating an alternative from the students’ and teaching perspectives. DE, for instance, challenges the suggestion “*dairy products*” in terms of the students’ knowledge of the word

“dairy” (i.e., *dairy’i bilemez/se/what if s/he does not know “dairy”*) and BA counter-challenges DE by making reference to shared past instructional events (Can Daşkın & Hatipoğlu, 2019; You, 2015) and explaining that “dairy” is a word that has been taught a lot to the students (i.e., *çok (.) öğrettik/we have taught it a lot*). This implies the reviewers’ epistemic right, responsibility and obligation (Heritage, 2013) to review the item with consideration of the context of learning as well as the abilities of the target group of students and thus, to create the interplay between teaching and testing. Since the issue concerns the students’ knowledge of a word, DE then backs off and uses a modified repeat (Stivers, 2005) of BA’s counter-evidence from past instructional events. This indexes her epistemic right to confirm an issue about the students’ L2 knowledge, asserting her epistemic authority and control as the teacher with access to the domain of teaching and the target students’ L2 abilities and knowledge (Heritage & Raymond, 2005; Stivers, 2005). Moreover, the reviewers’ shared role as teachers with equal epistemic access to the domain of teaching is manifested in BA’s use of an inclusive language for their teaching practice (i.e., *öğrettik/we have taught*).

Despite an orientation to the institutional role distribution of the participants, it is seen that they share the responsibility for reviewing the change in the problematic item. Not only the tester but also the reviewer DE reconsiders the validity of the new alternative through her confirmation check and the presentation of a challenge towards it to make sure once again that the change causes no problems that could go unnoticed. The participants follow a joint decision-making process in finalizing the item by seeking each others’ confirmation, playing the devil’s advocate, providing account and evidence in response to the presented challenges, together evaluating the alternative “*dairy products*” which emerges in the phase of RC in and through talk and displaying mutual agreement as well as shared understanding of the decided change. This is because the participants assume equal epistemic state and authority since they have equal epistemic access to the available resources (e.g., reading text) and share a common ground in terms of the testing principles for language test construction.

The participants’ equal active engagement in the phase of RC is also evident in their gaze behaviour. It is seen that throughout the interaction, they continuously look at their exam papers only sometimes directing their gazes at each other for a short time. DE, for instance,



scans through the text to check the availability of the suggested information in the reading text and BA silently reads a part of the original option (i.e., *°er designed by°*) while comparing the two alternatives to state her preference. However, this does not indicate disengagement; on the contrary, it signals that the participants actively think about the change made in the item in such a competing task-oriented activity (Goodwin, 1981; Rossano, 2013). Finally, EK's use of an inclusive language when seeking confirmation for the change in line 12 represents the participants' shared responsibility for the change process in the item.

Extract 2c presents the phase of RC in the reviewing interaction on the item problematized in extract 2a (i.e., phase PI) and revised in extract 2b (i.e., phase SC) to overcome the problem. It is to be remembered that in this extract, the two teachers YE and CI teaching at B1 level do the reviewing of a grammar item in a tense passage and provide their feedback to the tester SA. In the first segment of the extract (i.e., extract 2a, see Section 4.1), YE problematized the verb *“cover”* given in one of the blanks in the tense passage to be transformed into its past tense form, claiming that it was not taught in class. Addressing this problem, the IR group found an alternative for the problematic verb in the second segment of the extract (i.e., extract 2b, see Section 4.2) through a process of suggestion and decision making for change in the item. As a result, the verb *“cover”* was replaced with *“change”* in the phase of SC in extract 2b. However, the reviewing interaction for this particular item does not come to end when a decision is made on the change and the item is revised accordingly but continues for further review of the accepted change this time in the following segment of the extract (i.e., extract 2c) to make sure that the change is complete and the problem is solved and that it raises no other problems that could go unnoticed in the item. This extract was also presented at a data session. The analysis provided here is, thus, supported by the analysis carried out at this data session.

#### Extract 2c. Cover, Segment c

- 1 SA: [change san]ki (.) >anlam olarak da< [oturur]  
*“change” seems to fit in terms of meaning too*  
 + directs gaze to CI  
 + YE shifts gaze to SA
- 2 CI: [ °ş:ey° ] [er: ]  
*well*  
 + looks at her paper

- 3 SA: gibi di mi? [=sıkıntı (.) olmaz ]  
*like doesn't it? it wouldn't be a problem*
- 4 CI: [şey diyebiliriz ]  
*we can say this*  
 + YE turns gaze to CI
- 5 belki changed the inside of his taxi using it  
*maybe*  
 + reads from the target sentence with a pen in her hand  
 + SA shifts gaze to her paper  
 + YE turns gaze to her paper  
 + SA shifts gaze to CI and then back to her paper
- 6 falan gibi  
*something like that*  
 + shifts gaze to SA and YE
- 7 (0.3)
- 8 hani [changed with] it bi garip oluyo.  
*well "changed with it" sounds weird*
- 9 SA: [taxi ]  
 + reads from the target sentence on her paper  
 + positions hand on her paper in an attempt to note the suggested change
- 10 (0.2)
- 11 YE: huh tamam  
*okay*  
 + SA crosses out "with" on her paper  
 + shifts gaze to CI
- 12 SA: ta↑m[am with] yerine o zaman [using] diyelim  
*okay we shall say "using" instead of "with"*  
 + writes "using" in place of "with" on her paper  
 + YE turns gaze back to her paper
- 13 YE: [evet ] [evet ]  
*yes yes*  
 + CI nods head, turning gaze back to her paper
- 14 o (.) [cümleye uya- uyarlamak için e[vet =yani]  
*to adapt it for that sentence yes I mean*
- 15 SA: [tamam ] [tamam ]  
*okay okay*
- 16 YE: önemli olan orda (.) şeyi-  
*the important thing there is its thing*  
 + SA and CI shift gazes to YE
- 17 CI: doğru evet [daha iyi oluyo  
*yes right it becomes better*  
 + SA and CI turn gazes back to their papers
- 18 YE: [bilmeleri  
*for them to know*
- 19 SA: [tamam  
*okay*

20 (0.2)  
 21 huh-[huh  
 22 YE: [past tense'i[ni  
           **its past tense**  
 23 CI: [tamam  
           **okay**  
 24 (0.4)  
 25 SA: tamam  
           **okay**

In extract 2c, the change reviewing sequence is initiated by the tester SA's request for confirmation with regard to the appropriacy of the new verb "*change*" in the target sentence in terms of its meaning by directing her gaze towards CI, which elicits YE's gaze to SA. SA displays a stance of uncertainty about the issue while formulating her confirmation-seeking question via the hedges "*sanki/seems to*" and "*gibi/like*". Through her confirmation request, she reviews and challenges the change indirectly, drawing the reviewers' attention to a different issue that would possibly emerge as a result of the revision of the item (i.e., the grammatical accuracy of the sentence when the new verb "*change*" is inserted into it). Simultaneously with SA's turn, while CI is looking at the item on her exam paper, she attempts to say something as signalled by her use of the elongated hesitation markers "°ş:ey°/*well*" and "er:" in line 2 as well as a preface with an inclusive language in line 4 (*şey diyebiliriz/we can say this*) projecting an upcoming suggestion, at which point YE this time shifts her gaze to CI.

As SA and YE are both looking at CI, CI then suggests a change to make the verb "*change*" fit better in the sentence in line 5, addressing the same issue already raised by SA, indirectly claiming a problem in the changed item and thus, providing a non-confirming response to SA's question. In doing so, with a pen in her hand, CI reads the part of the target sentence following the new verb "*changed*" (i.e., *changed the inside of his taxi*) by inserting her suggestion "using it" in place of "*with it*" in the sentence and reading it out with an emphasis so that the other participants know what part of the sentence she is suggesting a change for. CI also downgrades her suggestion by using the hedges "*belki/maybe*" and "*gibi/something like*", and presenting it as an example for the group to consider (i.e., *falan gibi/something like that*). Just as she starts to read the sentence, the other participants SA and YE shift their gazes to their exam papers to follow CI and just as CI begins to articulate her suggestion, SA turns her gaze to CI and then

back to her paper. Just prior to the completion of CI's turn, CI shifts her gaze to SA and YE to elicit a response from them.

After a gap of brief silence as YE and SA are looking at the item on their exam papers and thinking of the suggested change, CI then provides an account for her suggestion in line 8 through her negative assessment that "changed with it" in the revised sentence sounds weird. Meanwhile, in overlap with CI's account, SA reads out the word "taxi" from the target sentence and spotting the trouble source "with" just following it, she positions her hand on her paper in an attempt to note the suggested change in place of the repairable, which shows her uptake of CI's suggestion. Then in line 11, YE displays agreement and a stance of acceptance of CI's suggestion through her use of the acknowledgement token "huh" and the acceptance token "tamam/okay" with a shift of gaze to CI. At this point, SA crosses out the trouble source "with" on her paper and then in line 12, she exhibits an upgraded agreement by using the acceptance token "ta↑mam/okay" with a rising intonation and offering a (re)formulation of CI's turn of suggestion (Heritage & Watson, 1979) with an inclusive expression of suggestion (i.e., *with yerine o zaman using diyelim/we shall say "using" instead of "with"*) while writing "using" in place of "with" on her paper. In this way, she not only displays her agreement with and understanding of the change but also establishes shared understanding in the group with regard to the suggested change and directs the trajectory of the interaction towards the final decision and the end of the reviewing interaction for this particular item. During SA's turn-in-progress, with a shift of gaze back to her paper, YE further acknowledges the new change through her use of the compliance token "evet/yes" twice in line 13 and hence, confirms SA's understanding of the change as correct while CI nods her head, turning her gaze back to her paper.

Following the suggestion and decision making sequence in the change reviewing process, YE initiates a non-minimal post-expansion sequence (Stivers, 2013) in line 14 to offer a candidate explanation for the suggestion, which makes the reason for the change more explicit and further promotes mutual understanding about it. Launching a self-initiated self-repair in the process of formulating her account (i.e., *uya- uyarlamak/to adapt*), she remarks and acknowledges that such a change is required to adapt the verb for the sentence. During YE's turn-in-progress, SA once again displays a stance of acceptance of

the change and shows agreement with YE through her use of the minimal expansion acceptance token “tamam/*okay*” twice in line 15, which also proposes sequence closure. However, YE then rushes through into another unit of talk in line 16 where she provides further clarification of the problem she has previously announced and brings the main issue to the forefront of the group’s attention by stating that the important point is for the students to know the past tense of the verb given in the item. In this way, she addresses the violated testing principle (i.e., “test what you teach”) and shows an orientation to her situated understanding of how a test should be.

As YE starts to reassert the main issue in the revision of the item, SA and CI shift their gazes to YE and then turning their gazes back to their papers, CI cuts off YE’s utterance in line 16 to pre-emptively convey an affiliative stance with YE and demonstrate her understanding of the issue raised by YE before YE completes her comment through her stance marker “doğru/*right*” and compliance token “evet/*yes*” as well as her positive assessment of the main change “daha iyi oluyo/*it becomes better*” in line 17. Meanwhile, SA continues to use the acceptance token “tamam/*okay*” repeatedly in lines 19 and 25 as well as the acknowledgement token “huh-huh” in line 21, which not only exhibits mutual agreement but also proposes sequence closure once again. CI, too, does the same in line 23 through her “tamam/*okay*” upon YE’s completion of her turn. As a result, the revision of the item is finalized and the reviewing interaction for this particular item ends.

Extract 2c demonstrates how the participants review the change that is already decided and noted in phase SC to test the past tense of a verb that is known by the students and how they finalize the reviewing of the item. The item change reviewing process begins with the recognition of a problem in the revised target sentence in the item as a result of the replacement of the problematic verb “cover” with “change” in phase SC. The participants realize that the new verb “change” does not fit grammatically well in the sentence. Carefully reviewing the revised sentence, SA’s confirmation request with a display of a stance of uncertainty (i.e., change sanki (.) >anlam olarak da< oturur gibi di mi?/*it is like “change” seems to fit in terms of meaning too, doesn’t it?*) and CI’s offer of a suggestion simultaneously with SA’s confirmation request (i.e., belki/*maybe* changed the inside of his taxi using it

falan gibi/something like) as well as her account for her suggestion (i.e., changed with it bi garip oluyo/sounds weird) all show an orientation to a problem that emerges in the item after the target sentence undergoes a change, shifting the focus of the interaction from “test what you teach” principle onto the grammaticality of the language used in the item. It becomes clear that the preposition “with” in the original sentence does not collocate with the new verb “change” and is grammatically inaccurate.

Consequently, SA and CI’s indirect challenging of the revised sentence determines the rest of the interaction in phase RC in extract 2c and the sentence eventually undergoes another change to prevent other problems that would emerge as a result of the major change made in the item to comply with the testing principles. This indicates that language test IR interactions require continuous reviewing and editing since a major change (i.e., replacing the verb in the item) might lead to other changes (i.e., changing the preposition of the verb expression) to make the revision of the item complete. This could happen either within the phase of SC before the decided change is fully acknowledged and noted on the exam paper as seen in extracts 8b and 10b or in a separate phase as in this extract where the participants review the already noted change before proceeding with the other items in the test.

The analysis has shown that a new suggestion-making sequence takes place in the phase of RC following the same structural organization that appears in the phase of SC. To make the decided verb “change” fit better in the target sentence in the item and thus, to make the sentence grammatically accurate, CI suggests replacing “with” in the original sentence by “using”, which is then met with the group’s uptake and put to use. What seems to play an important role in the acceptance of CI’s suggestion is her account-giving, which makes it clear that “changed” does not collocate with the preposition “with”, causing the verb expression “changed with it” to sound weird. Since CI offers her suggestion without orienting to SA’s confirmation request and explicitly stating the problem in the sentence, it is only after her account following her suggestion that the other co-participants show an explicit agreement and display of understanding. Before providing her account, the suggestion receives no immediate response and both SA and YE orient to the test material, trying to understand the suggested change. Therefore, as it was found in the phase of SC, the suggestion that is justified based on evidence is more likely to be accepted.

Regarding the act of suggesting a change, as in the phase of SC, CI downgrades her suggestion by using hedges such as “belki/*maybe*” and presents it just as an example alternative that can be either accepted or rejected by the group (i.e., *falan gibi/something like that*). In this way, CI does not implicate epistemic priority as the first provider of the suggestion (Heritage & Raymond, 2005) but instead invites collaborative thought on the issue and solicits mutual agreement, sharing the responsibility for the decision. Thus, CI might not want to block the other participants’ contributions by being too assertive about her suggestion but rather pave the way for other alternatives and for a careful thought and decision-making on the sentence. This is because the participants have equal epistemic access to the referent situation (e.g., the test material) and they are in K+ position in terms of their English language expertise as English teachers, who can equally judge the grammaticality of the sentence and edit it where necessary (Heritage, 2013). On the other hand, CI’s subsequent display of certainty with her account for her suggestion through a stronger evaluative stance (i.e., *changed with it bi garip oluyo/sounds weird*) leads to the acknowledgement of the problem and the acceptance of her suggestion by the group. Furthermore, as a commonly used interactional resource in the phase of SC, CI applies a preface (i.e., *şey diyebiliriz/we can say this*) which initiates the suggestion-making sequence by projecting the announcement of her upcoming suggestion and draws the other participants’ attention to it.

The participants employ some interactional resources to display mutual agreement and to ensure shared understanding about the change in the item so that a joint decision-making process can be followed. This is achieved first by SA’s (re)formulation of the suggested change (Heritage & Watson, 1979) while noting it on her exam paper by referring to the trouble source in the sentence and repeating the suggestion that is to replace it (i.e., *with yerine o zaman using diyelim/we shall say “using” instead of “with”*). In this way, she not only shows her uptake of CI’s suggestion but also ensures correct and shared understanding of the change so that it can be noted correctly on the exam paper.

Additionally, the non-minimal post-expansion sequence plays an important role in establishing intersubjectivity in the IR group. Even though the participants’ display of agreement with regard to the suggestion does not project an expansion, YE initiates a non-

minimal post-expansion sequence for further elaboration and for final comments about the item (Stivers, 2013). To illustrate, although YE initially shows agreement with the suggested change through her acceptance and compliance tokens, she then demonstrates her understanding of the reason for such a change by offering a candidate explanation for it (i.e., o (.) cümleye uya- uyarlamak için/to adapt it for that sentence). She not only elaborates on the new change that is processed in phase RC but also on the main issue that she raised earlier in phase PI and that led to the change process in the item by referring to the violated testing principle and reminding that what actually matters is the students' knowledge of the past tense of the verb that is tested (i.e., önemli olan orda (.) şeyi- bilmeleri past tense'ini/the important thing there is its thing for them to know its past tense). By bringing to the forefront the underlying reason for the change process in the item and clarifying it, she makes sure that the main issue is not overlooked as they shift their focus to the grammaticality of the sentence and ensures shared understanding this time with regard to their primary aim in revising the item. Furthermore, CI exhibits mutual agreement with YE and makes her stance clearer through her pre-emptive positive assessment of the main change addressed by YE (i.e., daha iyi oluyo/it becomes better). The participants' display of agreement and understanding with regard to the changes made in the item via such interactional resources overtly confirms the revised item and eventually leads to the final decision, ending the reviewing interaction for the item.

Considering the institutional role distribution of the participants in extract 2c, there are indications of tester, reviewer and teacher roles and responsibilities in this IR interaction. To begin with the tester roles, it is seen that as the writer of the test item who is accountable for the test, the tester SA often seeks final confirmation and mutual agreement with respect to the revised item before finalizing it. She does it first through her confirmation request when reviewing the item in terms of the grammaticality of the modified target sentence where the original verb "cover" is replaced with "change" and initiating the change reviewing process (i.e., change sanki (.) >anlam olarak da< oturur gibi di mi? =sıkıntı (.) olmaz/it is like "change" seems to fit in terms of meaning too doesn't it? it wouldn't be a problem). Secondly, she ensures correct and mutual understanding of the second change that the group decides on in phase RC by explicitly articulating the kind of change she is making while noting it down on



her exam paper through a (re)formulation of CI's turn of suggestion. In doing this, she refers to the trouble source that is to be replaced in the target sentence and restates the suggested change (i.e., with *yerine o zaman using diyelim/we shall say "using" instead of "with"*). The tester's such an act of (re)formulating the decided change functions as a final check for the item, orienting to her role as the tester who is accountable for the test. In this way, as the tester who is responsible for producing the final product, she avoids an incorrect noting of the change on the test material. In addition, SA's non-verbal behaviour of noting down the second change on the test material orients to her responsibility and obligation to finalize the test item and make it ready for administration. Lastly, as the tester, SA monitors the item change reviewing process by proposing sequence closure through her frequent use of the minimal expansion acceptance token "*tamam/okay*" upon the group's agreement with and confirmation of the second change and thus, directing the trajectory of the interaction towards the final decision so that the IR interaction can proceed for the other items on the test material.

As for the teacher roles, the reviewer YE displays an orientation to her role as the teacher by making reference to the violated testing principle from the students' perspective (i.e., *bilmeleri/for them to know*). Using the third-person plural pronoun to refer to the students, she, thus, clarifies the previously announced problem in terms of their knowledge of the verb given in the blank to be transformed into its past tense form. This implies the reviewers' epistemic right, responsibility and obligation (Heritage, 2013) to review the item with consideration of the abilities of the target group of students. YE's further clarification of the problem she announced in phase PI at the same time orients to her role as the problematizer of the item as the reviewer BA does in extract 1b. This signals her responsibility to remind the IR group of the main issue and to establish common ground on the underlying reason for the revision of the item at the point when the participants shift their focus to the grammatical accuracy of the revised sentence while reviewing it. Hence, as the problematizer of the item, she makes sure that the problem is addressed and is clearly understood by the group.

Despite an orientation to the institutional role distribution of the participants, it is seen that they share the responsibility for reviewing the change in the problematic item. In this particular extract, the participants all have a contribution to make to the reviewing process.

While SA raises the problem in the grammatical structure of the revised sentence through her confirmation request, CI suggests a change to overcome this problem (i.e., replacing the preposition “*with*” in the sentence by “*using*”) and YE offers a candidate explanation for the suggested change and then provides a further clarification of the main issue (i.e., students’ knowledge of the verb given in the blank) in an attempt to achieve intersubjectivity in the group and make sure that the changes are built on a solid ground. In employing all these actions, the participants also seek each others’ confirmation as SA does, for instance, through her confirmation request and as CI does by downgrading her suggestion. Thus, all the three participants actively and critically think about the revised item in and through talk and follow a joint decision-making process.

The participants’ equal active engagement is also evident in their gaze behaviour as they continuously look at their exam papers, constantly reviewing the change and checking the appropriacy of the new suggestion (Goodwin, 1981; Rossano, 2013). This is because the participants assume equal epistemic state and authority since they have equal epistemic access to the available resources and are in K+ position in terms of their English language expertise (Heritage, 2013). Finally, the participants’ use of an inclusive language throughout the interaction (i.e., *diyebiliriz/we can say*; *diyelim/we shall say*) represents their shared responsibility for the change process in the item and their joint decision-making and shows that the change is the product of the whole group.

Extract 3c presents the phase of RC in the reviewing interaction on the item problematized in extract 3a (i.e., phase PI) and revised in extract 3b (i.e., phase SC) to overcome the problem. It is to be remembered that in this extract, the two teachers BA and DE teaching at B1 plus level do the reviewing of a grammar item in a multiple-choice cloze test and provide their feedback to the tester ME. This item tests the students’ knowledge of the conjunctions they studied in class. In the first segment of the extract (i.e., extract 3a, see Section 4.1), the reviewers problematized the item, claiming that the distractor “*therefore*” could also be possible as a correct answer besides the expected answer “*furthermore*”. Addressing this problem, the IR group found an alternative for the other possible correct answer “*therefore*” in the second segment of the extract (i.e., extract 3b, see Section 4.2) through a suggestion and decision making process. As a result, the problematic distractor “*therefore*” was replaced with “*whereas*” in the phase of SC in extract 3b. However, the reviewing interaction for this particular item does not come to end when a decision is made on the



- 16 BA: ta↑bi  
+ nods head  
**of course**
- 17 ME: °veya [noktalı] [virgül°  
**or semicolon**  
+ DE shifts gaze to her paper
- 18 DE: [huh ]
- 19 BA: [ba- ara bağlaç olma[sı lazım  
**it should be a coordinating conjunction**  
+ DE nods head
- 20 ME: [aynen öyle  
**that's right**
- 21 (0.2)
- 22 DE: huh-↑huh
- 23 BA: tamam  
**okay**
- 24 ME: °güzel oldu bu da°  
**this has also become nice**

\* On the exam paper, the word is actually written as “sheet” but is produced as “street” by BA.

As DE and BA complete their noting of “*whereas*” in place of the problematic distractor “*therefore*” on their exam papers and ME is looking at her exam paper having already made the change on her paper in phase SC, extract 3c begins with DE’s initiation of the item change reviewing process by translating the new distractor “*whereas*” into Turkish (i.e., *halbuki*) and checking its meaning within the context given in the cloze text and then launching a confirmation check with an initial pitch rise (i.e., ↑*oldu mu?/is it all right?*) and requesting the group to review the change made in the item (i.e., *bi bakın/have a look at it*). Just prior to the completion of her turn, DE directs her gaze towards BA’s exam paper. BA takes over in line 4 to provide a repetitional response (Heritage & Raymond, 2012) that overtly confirms the change in alignment with the terms and constraints of DE’s question (i.e., *oldu/it is all right*).

Playing the devil’s advocate, DE then employs another confirmation check in line 6 by redesigning her initial confirmation-seeking question to clarify her point (i.e., *olmaz yani di mi?/I mean it would not be possible would it?*) with a shift of gaze back to her paper. This time she requests and checks the group’s confirmation particularly with regard to the possibility of the new distractor as a correct answer, which is the main

concern in revising the item. She, thus, orients to the violated testing principle that each item should have only one correct answer (Heaton, 1990; Hatipoğlu, 2009), making sure that the identified problem in the item is solved. In overlap with DE's question, BA starts to review the change following an elongated hesitation marker by reading out the target sentence between lines 7 and 12 and inserting the new conjunction "*whereas*" into the sentence with an emphasis and a rising intonation on its final syllable while DE and ME look at their exam papers. Immediately following her reading of the sentence, BA rushes to provide a confirming response to DE through her use of the marked confirmation "*tabi/of course*" (Stivers, 2011) in line 12, which further acknowledges the appropriacy of the new option as a distractor.

During BA's turn-in-progress as she is reading out the sentence, DE and ME give an account for the incorrectness of the new distractor as an answer from different perspectives. While DE presents evidence in terms of punctuation rules, ME considers it in terms of the content of the target sentence. Shifting her gaze to BA's exam paper and pointing at the relevant part of her paper with a pen, DE explains that "*whereas*" would not appear after a full stop as in the case of the target sentence, which makes the new distractor unlikely as a correct answer and solves the problem. Simultaneously with DE and BA's turns, shifting her gaze to BA in line 14, ME points out that different from the meaning expressed by "*whereas*", no contrast is implied in the target sentence and hence, offers a confirming response to DE's question indirectly through account-giving, which is made conditionally relevant by DE. Through the end of DE's account in line 15, DE and ME turn their gazes back to their papers and then DE shifts her gaze to BA, eliciting a response from her. As a result, BA displays an upgraded agreement with DE in line 16 by articulating the marked confirmation "*ta↑bi/of course*" (Stivers, 2011) with a rising intonation accompanied by her head nod.

ME then provides a supportive and complementary explanation with a soft tone of voice in line 17 in relation to DE's account about the punctuation used with "*whereas*" by adding semicolon to the full stop mentioned by DE as the incorrect punctuation for "*whereas*", at which point DE shifts her gaze to her paper. During ME's turn-in-progress, DE shows agreement through the acknowledgement token "*huh*" in line 18 and overlapping ME's turn at a turn-final position in line 19, BA brings in additional knowledge that the sentence

requires a coordinating conjunction, which further functions as complementary to DE's explanation, supporting and endorsing her stance. In this way, through their accounts, the participants all bring professional knowledge to the interaction and show expertise as language teachers while reviewing the change and presenting evidence that proves the unlikelihood of the new distractor "*whereas*" as a correct answer in response to DE's confirmation-seeking question.

DE exhibits agreement with BA by nodding her head through the end of BA's turn-at-talk. Overlapping BA's turn at a turn-final position, ME conveys an affiliative stance and displays a strong alignment with BA in line 20 via her use of the agreement token "*aynen öyle/that's right*". After a gap of brief silence, DE once again shows agreement with the group in line 22 by using the acknowledgement token "*huh-↑huh*" with a rising intonation. BA then displays a stance of acceptance of the change and proposes sequence closure through her use of the minimal expansion acceptance token "*tamam/okay*" in line 23. Finally, ME offers a positive assessment of the change with a soft tone of voice in line 24, which further confirms the revised item and ends the reviewing interaction for this particular item. As a result, the participants all acknowledge the change made in the item and establish mutual understanding and agreement in the IR group with respect to it.

Extract 3c demonstrates how the participants review the change that is already decided in the phase of SC to make the distractor option "*therefore*" unlikely as a correct answer and how they finalize the reviewing of the item. This phase can be characterized by such interactional actions as confirmation checks and account-giving in an attempt to make sure that the change addresses the violated testing principle and to achieve shared understanding with regard to it before ending the reviewing interaction for this particular item. As the suggestion-maker, the reviewer DE is active in monitoring the review process for the change made in the item. She launches two confirmation checks (i.e., *↑oldu mu? bi bakın/is it all right? have a look at it; olmaz yani di mi?/I mean it would not be possible would it?*) whereby she seeks and checks the group's confirmation for the revised item. Meanwhile, she herself already begins to review the change by translating the new distractor "*whereas*" into Turkish (i.e., *halbuki*) to check its meaning within the context given in the target sentence and providing account and evidence for it (i.e., *noktadan sonra da olmaz whereas/"whereas"*

*also wouldn't be okay after full stop*) before waiting for the other participants' responses to her second confirmation request.

DE's confirmation checks play an important role in inviting the group to re-evaluate the change as BA starts to read the target sentence by inserting the new distractor "*whereas*" into it and ME keeps her gaze at her paper for some time before making a comment on the change in line 14. By redesigning her initial confirmation-seeking question in line 6 (i.e., *olmaz yani di mi?/I mean it would not be possible would it?*) and requesting the group's confirmation with regard to the incorrectness of the new distractor "*whereas*" as the main concern in revising the item, DE also orients to the violated testing principle that each item should have only one correct answer (Heaton, 1990; Hatipoğlu, 2009) and thus, directs the group to consider if the problem identified in the first phase of the IR interaction is solved. As a result, DE's confirmation requests initiate a discussion about the appropriacy of "*whereas*" as a distractor and lead to a collaborative process of reviewing the change carefully. In the end, the change is eventually acknowledged and mutual understanding is developed in the group in and through the talk-in-interaction.

Secondly, the interactional act of account-giving is a predominant feature of the interaction in extract 3c. It is seen that the participants provide account for the change in the item collaboratively from different perspectives and confirm the revised item based on evidence. To illustrate, first ME evaluates the change in terms of the content of the target sentence (i.e., *zıtlık yok/there is no contrast*) and justifies the unlikelihood of the new distractor "*whereas*" as a correct answer on the grounds that the sentence does not express a meaning of contrast as implied by "*whereas*", proving that it solves the problem of "more than one correct answer" identified in the item. On the other hand, DE examines the change in terms of the punctuation rules for "*whereas*" (i.e., *noktadan sonra da olmaz whereas/"whereas" also wouldn't be okay after full stop*), presenting the full stop used in the target sentence as evidence that makes the distractor incorrect as an answer. DE's account is then extended by ME as she adds semicolon to the full stop mentioned by DE as the incorrect punctuation for "*whereas*" (i.e., *°veya noktalı virgül°/or semicolon*). Finally, BA offers an explanation in terms of the type of conjunction that is required in the target sentence (i.e., *ara bağlaç olması*

*lazım/it should be a coordinating conjunction*) and further proves the appropriacy of “*whereas*” as a distractor based on the assumed shared knowledge that it does not fall into the group of coordinating conjunctions as necessitated by the sentence. Such a collaborative process of account-giving demonstrates that the participants all bring their professional expertise to the interaction and show their subject knowledge as language teachers while reviewing and justifying the change in the problematic distractor in terms of its unlikelihood as a correct answer and hence, the violated testing principle that each item should have only one correct answer (Heaton, 1990; Hatipoğlu, 2009). Consequently, their accounts lead to the acknowledgement and approval of the change as well as the establishment of shared understanding and mutual agreement on the revised item in the IR group.

The participants do not just use an agreement token but employ some interactional resources to display mutual agreement so that the change can be overtly confirmed and a joint decision-making process can be followed. To begin with, BA’s repetitional response (Heritage & Raymond, 2012) in line 4 and her use of the marked confirmation “*tabi/of course*” (Stivers, 2011) twice in lines 12 and 16 not only exhibit her agreement with the change and with the other participants’ accounts but also display a stance of certainty towards the revised item, which results in the acceptance of the new distractor “*whereas*” noted on the exam paper. Secondly, ME and BA’s supportive and complementary explanations in relation to DE’s account in lines 17 and 19 not only show their agreement with DE but also bring in additional knowledge that further confirms the change in the item. Upon DE’s account with regard to the punctuation used in the target sentence (i.e., full stop), ME makes reference to another punctuation that does not allow the use of “*whereas*” in a sentence (i.e., semicolon) while BA mentions the type of conjunction that is required in the target sentence (i.e., coordinating conjunction) and that does not include “*whereas*”. In this way, the participants collaboratively provide an account for the unlikelihood of “*whereas*” as a correct answer and ensure shared understanding in the group. By means of such interactional resources, the participants also claim their epistemic right to confirm an assertion, indexing epistemic authority and control over the matter at hand since they all have equal epistemic access to the test material and are in K+ position in terms of their English language expertise, which in this case includes their knowledge of conjunctions (Heritage & Raymond, 2005; Stivers, 2005). Finally, ME’s positive assessment of the



change (i.e., °güzel oldu bu da°/*this has also become nice*) is another interactional resource used in the extract to confirm the change and display mutual agreement, which eventually leads to the final decision and ends the reviewing interaction for the item.

This extract demonstrates that the participants share the responsibility for reviewing the change in the item and are equally engaged in phase RC. They all review the change carefully by, for instance, reading out the target sentence with the insertion of the new distractor “*whereas*” into it, translating it into Turkish, providing account and evidence for its unlikelihood as a correct answer, and bringing their professional expertise to the interaction to explain why “*whereas*” becomes an incorrect answer as a distractor. In addition, a joint reviewing process is encouraged by DE’s confirmation-seeking questions and her invitation to recheck the revised item (i.e., *bi bakın/have a look at it*). In doing this, she does not just wait for the other participants to do the reviewing of the change but she herself starts to evaluate it simultaneously and even provides the first account before the other participants respond to her confirmation request. Thus, all the three participants actively and critically think about the revised item in and through talk, follow a joint decision-making process, and eventually reach a shared understanding about the change. Their active engagement is also evident in their gaze behaviour as they continuously look at their exam papers, constantly reviewing the change and checking the appropriacy of the new distractor (Goodwin, 1981; Rossano, 2013). This is because the participants assume equal epistemic authority since they have equal epistemic access to the resources and are in K+ position in terms of their English language expertise (Heritage, 2013).

Extract 4c presents the phase of RC in the reviewing interaction on the item problematized in extract 4a (i.e., phase PI) and revised in extract 4b (i.e., phase SC) to overcome the problem. It is to be remembered that in this extract, the two teachers DU and GE teaching at B1 level do the reviewing of a multiple-choice reading comprehension item and provide their feedback to the tester SA. In the first segment of the extract (i.e., extract 4a, see Section 4.1), DU problematized the distractor option “b” (i.e., *customers do shopping very fast*) and claimed that it could also be possible as a correct option considering the stem “*when Muzak soundscapes are used*” and the information in the reading text which was about a company named “Muzak” making specific music appropriate for different businesses. Addressing this problem, the option in question underwent change in the second segment of the extract (i.e.,

extract 4b, see Section 4.2) as the IR group redesigned it through a process of suggestion and decision making for change in the item. As a result, the problematic distractor option “customers do shopping very fast” was replaced with “customers prefer to buy expensive products” in the phase of SC in extract 4b. However, the reviewing interaction for this particular item does not come to end when a decision is made on the change and the item is revised accordingly but continues for further review of the accepted change this time in the following segment of the extract (i.e., extract 4c) to make sure that the change is complete and the problem is solved and that it raises no other problems that could go unnoticed in the item.

#### Extract 4c. Muzak Soundscapes, Segment c

- 1 SA: <sup>1#</sup>bu: >muhtemelen biraz< uzun olacak >diğerlerine göre<  
+ DU and GE look at their exam papers + DU directs gaze to SA's paper
- 2 ama  
**but this will probably be a little longer compared to the others**
- 3 (4.5) <sup>#1</sup>  
#1 (lines 1-3) SA notes the new option on her exam paper and upon its completion, she then produces thinking face by putting her finger on her mouth while DU turns her gaze back to her paper
- 4 GE: buy [dersin °hani °]  
**well say "buy"**  
+ shifts gaze to SA and then to DU
- 5 DU: [customers ] ↑bu:y=
- 6 GE: ev[et pre[fer'ü yazma ]  
**yes don't write "prefer"**  
+ shifts gaze to SA and then back to her paper
- 7 DU: =[expen [sive products de]  
**say**  
+ shifts gaze to the new option written on SA's paper
- 8 SA: [EVET prefer'ü ] atıyorum [doğru  
**yes let me omit "prefer" right**  
+ crosses out the word “prefer” noted on her exam paper
- 9 DU: [prefer'ü at  
**omit "prefer"**  
+ turns gaze back to her paper
- 10 SA: doğru (0.2) tamam  
**right okay**
- 11 (1.2)



- + turns over the page to look at the reading text
- 30 SA: [hu:h]
- 31 DU: çünkü olabilecek tek (.) er şey (.)
- 32 [bakacak ] inceli[cek ]  
***because it is the only thing that can be possible,  
s/he will look and examine***
- 33 SA: [ °huh-huh ° ] [evet]  
***yes***

As SA notes the new option “*customers prefer to buy expensive products*” on her exam paper and as DU and GE are looking at their exam papers, extract 4c begins with SA’s challenging of the change she is making on her paper with a turn-final “*but*” by pointing out that the new option would end up being longer than the other options in the item. She, thus, refers to the violation of a testing principle about the format of options in multiple-choice test items, which states that “options should be similar in structure” (Fulcher, 2010, p.172) and more specifically that the length of options should be kept about equal (Haladyna & Rodriguez, 2013). During SA’s turn-in-progress, DU directs her gaze towards SA’s exam paper to check the length of the option as SA is writing it on her paper. SA’s challenging is followed by a lapse of 4.5 seconds when SA completes her noting of the change and then produces thinking face by putting her finger on her mouth and DU turns her gaze back to her paper having checked the length of the option once it is fully written on paper.

As all the participants are looking at the item on their exam papers, thinking about the issue raised by SA, GE and DU come up with the same suggestion and offer it simultaneously in lines 4 and 5. To make the option shorter, they both suggest leaving out “*prefer*” in the option and simply stating “*customers buy expensive products*”. While making her suggestion, GE shifts her gaze to SA and then to DU as DU overlaps to propose the same solution in lines 5 and 7 by reading out the option with “*prefer*” omitted with an initial pitch rise and an elongation on the word “↑bu:y” and shifting her gaze to the new option written on SA’s paper. In line 6, GE shows agreement with DU through the compliance token “*evet/yes*” and overlapping DU’s turn-in-progress, she extends her turn to clarify her suggestion by telling SA not to write “*prefer*” with a shift of gaze to her and then back to her paper.

SA displays agreement with the suggestion in line 8 by articulating the compliance token “EVET/*yes*” at a higher volume and demonstrates her understanding of the suggestion by offering a (re)formulation of GE’s suggestion (Heritage & Watson, 1979) and explicitly indicating the kind of change she is to make in the option (i.e., *prefer’ü atıyım/let me omit “prefer”*) followed by the stance marker “doğru/*right*”, which further acknowledges the suggested change. Her turn is embodied as she crosses out the word “*prefer*” noted on her paper, which shows her uptake of the offered change. Overlapping SA’s turn at a turn-final position, DU also clarifies her suggestion in line 9 by using a modified confirming repeat (Kim, 2002; Stivers, 2005) of SA’s turn and telling SA to leave out “*prefer*” with a shift of gaze back to her paper. The suggestion-makers also adopt a stance of certainty towards their suggestions by uttering them in declaratively-formatted turns and imperative forms (e.g., *dersin/say, de/say, yazma/don’t write, at/omit*). SA shows agreement and an acceptance of the new change once again in line 10 through the stance marker “doğru/*right*” and the minimal expansion acceptance token “*tamam/okay*”.

After a gap of silence, another item change reviewing sequence is initiated by SA in line 12 as all the participants are looking at their exam papers. Directing it to the teachers with the address term “*hocam/miss*” in turn-final position accompanied by her shift of gaze to GE, this time SA employs a confirmation check with respect to the incorrectness of the new distractor as an answer, which is the main concern in revising the item. She, thus, makes a reference to the violated testing principle that each item should have only one correct answer (Heaton, 1990; Hatipoğlu, 2009), making sure that the identified problem in the item is solved. SA, then, rushes to provide an account in line 13 for the appropriateness of the change in terms of making the distractor incorrect as an answer. Shifting her gaze to DU, she explains that the information in the new option is not available in the reading text.

Just as SA begins to give an account, the other participants provide a confirming response to her confirmation-seeking question as GE nods her head and DU uses the compliance token “*evet/yes*” twice consecutively in line 14 followed by her repetitional response (Heritage & Raymond, 2012) with a shift of gaze to SA. This overtly confirms that the change makes the distractor option wrong in alignment with the design of SA’s question. Overlapping SA and DU’s turns at a turn-final position, GE further orients to SA’s confirmation request in

line 15 and explicitly acknowledges that the distractor option becomes wrong through a repetitional response (i.e., *yanlış/wrong*) (Heritage & Raymond, 2012). GE and DU then display an orientation to SA's account in lines 16 and 17 and show agreement on the unavailability of the new information in the reading text by employing a modified confirming repeat (Kim, 2002; Stivers, 2005) of SA's turn of account (i.e., *yok/no; öyle bi bilgi yok/there is no such information*) with GE's shift of gaze to SA. During GE's turn-in-progress, SA shifts her gaze to GE by nodding her head and then uses the minimal expansion acceptance token "*tamam/okay*" in line 18, which exhibits her agreement with the other participants. Upon the group's acknowledgement of the change, they all shift their gazes back to their papers and SA proposes sequence closure by repeating the acceptance token "*tamam/okay*" in line 19.

DU initiates a non-minimal post-expansion sequence in line 20 for further elaboration of and account-giving for the change in the item (Stivers, 2013). Different from SA's account, DU presents more concrete evidence from the reading text by indicating that the text does not mention anything about the expensiveness or cheapness of the products that are bought as stated in the new distractor option. She, thus, supports and endorses SA's justification with regard to the unlikelihood of the option as a correct answer and shows that the problem of "more than one correct answer" identified in the item is solved. This is met with the other participants' agreement as they use the compliance token "*evet/yes*" in lines 21 and 23 with SA's shift of gaze to DU. SA also articulates the acknowledgement token "*huh-huh*" in line 22.

Turning her gaze back to her paper, SA then rushes to offer an additional explanation for DU's account in lines 24 and 25 with a rising intonation on the second syllable of the comparative determiner "*da ↑ ha/more*". By making reference to the relevant part of the reading text, she explains that the Muzak soundscapes make people do more shopping but not necessarily expensive ones. Hence, she makes it clearer that in the text, the expensiveness of the products is not the point in question, which proves the unlikelihood of the new distractor as a correct answer. This is received with the other participants' display of agreement as DU uses the compliance token "*evet/yes*" in overlap with SA's turn-in-progress in line 26 and GE uses the agreement token "*°aynen°/exactly*" with a soft tone of voice in line 29 by turning over the page to look at the reading text.

After a gap of brief silence, DU makes a final comment about the plausibility of the option and the difficulty level of the revised item from the students' perspectives. According to her, the change makes the item more difficult as the students will have to carefully read and examine the text, which fulfils other testing principles for multiple-choice item construction (i.e., plausibility of options, difficulty level of the item). During DU's turn-at-talk, SA displays agreement with DU through her use of an acknowledgement token twice in lines 30 and 33 followed by the compliance token "evet/yes". As a result, the revised item is approved and mutual understanding is achieved in the group, ending the reviewing interaction for this particular item.

Extract 4c demonstrates how the participants review the change that is already decided in the phase of SC to make the distractor option "*customers do shopping very fast*" unlikely as a correct answer and how they finalize the reviewing of the item. The interactional actions that play an important role in initiating and monitoring the item change reviewing process in this extract are the tester SA's challenging of the change and her confirmation check in an attempt to make sure that the change fulfils not only the violated testing principle but also other testing principles for multiple-choice item construction, causing no additional problems and to achieve shared understanding with regard to it before ending the reviewing interaction for this particular item. To begin with the former, through the act of challenging the change (i.e., bu: >muhtemelen biraz< uzun olacak >diğerlerine göre< ama/*but this will probably be a little longer compared to the others*), an orientation is displayed to the violation of a different testing principle for multiple-choice test items which was caused by the new option and which could not be foreseen until it was fully noted on the exam paper. The testing principle in question is not about the content of the option but about its format, which states that "options should be similar in structure" (Fulcher, 2010, p.172) and more specifically that the length of options should be kept about equal (Haladyna & Rodriguez, 2013). This is to "avoid giving clues to the right answer" (Haladyna & Rodriguez, 2013, p.91) since "if all else fails in a multiple-choice test, students will select the longest option" (Fulcher, 2010, p.172). In this case, a longer distractor might mislead the students by causing them to regard it as the correct answer.

Consequently, the act of challenging leads to the recognition of a problem that emerges in the item after the problematic distractor undergoes a change, shifting the focus of the

interaction from the content of the option and “only one correct answer” principle as the main concern onto its format. Based on this, the matter at hand is displayed and regarded as assumed knowledge with respect to the testing principles adopted for multiple-choice item construction within the institutional context in question since it is not further discussed and the group directly orients to a search for a change in the option to overcome the problem raised by SA, which signals their acceptance of the presented challenge. Therefore, the act of challenging determines the rest of the sequence as the option undergoes another change (i.e., omission of the verb “*prefer*”) to prevent other problems that could arise as a result of the major change made in the item and hence, to comply with the testing principles. This indicates that language test IR interactions require continuous reviewing and editing since a major change (i.e., changing the content of the distractor option) might lead to other changes (i.e., shortening the option) to make the revision of the item complete. In this respect, this phase can be considered to be a combination of phases one and two where another problematizing and change processes are carried out together.

As for the tester SA’s final confirmation check (i.e., *yanlış oluyo di mi? hocam/it becomes wrong doesn’t it miss?*), it initiates a new item change reviewing sequence in the phase of RC in extract 4c where SA seeks and checks the group’s confirmation for the revised distractor option with respect to its incorrectness as an answer as the main issue. Her confirmation request serves several functions in the change reviewing interaction. Firstly, it invites the group to re-evaluate the change and hence, leads to a collaborative process of reviewing the change, initiating a discussion about the appropriacy of the new option. Secondly, shifting the focus of the interaction from the format of the option back to its content, it makes a reference to the violated testing principle that each item should have only one correct answer (Heaton, 1990; Hatipoğlu, 2009) and thus, directs the group to consider if the problem identified in the first phase of the IR interaction is solved. In the end, the change is eventually acknowledged and mutual understanding is developed in the group in and through the talk-in-interaction.

The analysis has shown that a new suggestion-making sequence takes place in the phase of RC following the same structural organization that appears in the phase of SC. To make the revised distractor option shorter so that the options can be similar in length, the reviewers suggest omitting “*prefer*” from the new option, which is then quickly met with the tester’s uptake and put to use. What seems to play an important role in the acceptance of the



suggestion is the reviewers' display of a stance of certainty towards it by presenting it in declaratively-formatted turns and imperative forms (e.g., *dersin/say*, *de/say*, *yazma/don't write*, *at/omit*). In addition, the suggestion-making sequence shows that the participants all make an attempt to ensure shared understanding about the suggestion by extending their turns and making further clarification. As the suggestion-makers, although GE and DU initially tell SA what to include in the option (i.e., *buy*), they then instruct her what not to write (i.e., *prefer*) to clarify which word SA should leave out as the next change, bringing the unnecessary word "*prefer*" to the forefront of the group's attention.

The interactional act of account-giving is a predominant feature of the interaction in extract 4c. It is seen that the participants provide account for the change in the item collaboratively from different perspectives and confirm the revised item based on evidence. In response to the tester's confirmation check, account and evidence that were already given in the previous phase are recalled, re-checked in the reading text and provided again in this phase. To illustrate, the participants first evaluate the appropriateness of the new distractor in terms of its unlikelihood as a correct answer by presenting evidence from the reading text. In this way, they ensure the testing principle of "only one correct answer" (Heaton, 1990; Hatipoğlu, 2009). Their accounts also proceed from general (i.e., *çünkü böyle bi bilgi yok genelde/because there is no such information in general*) to more specific (e.g., *pahalı ya da ucuzluğuna dair bişey demiyo zaten/after all it does not say anything about its expensiveness or cheapness*) as they carefully re-examine the reading text and bring in concrete evidence to back up their confirmation that the text makes no reference to the information in the new option and that it has no possibility of being a correct answer as a distractor. They not only limit their account to the absence of the new information in the text but also check to see in particular whether the text includes anything about the expensiveness or cheapness of the products people buy as stated in the revised distractor. They eventually find out that it mentions only doing more shopping but not necessarily expensive ones. In doing this, while DU refers to what is not said in the text, SA refers to what is said in the text about the issue, evaluating the information in the text in relation to the new option from different dimensions.

Secondly, when making her final comment about the item, DU provides an account for the change by evaluating the revised item from the students' perspectives (i.e., *en zor zoru- (.) soru da bu olur bence çünkü olabilecek tek (.) er şey (.) bakacak incelicek/I think this would also be the most difficult question because it is the only thing that can be possible, s/he will look and examine*). By explaining that the item is a difficult one, requiring the students to carefully examine the text, DU displays an orientation to the testing principles about the plausibility of options (i.e., options should be reasonably attractive and plausible) and the difficulty level of the item (i.e., the item should be at a level appropriate to the proficiency level of the test-takers), making sure that the item fulfils the other testing principles too. Furthermore, even though the participants' display of agreement with regard to the revised item through their acceptance and compliance tokens does not project an expansion, the act of account-giving continues in a non-minimal post-expansion sequence (Stivers, 2013) through a collaborative process, which plays an important role in establishing mutual understanding about the revised item before fully confirming and finalizing it.

The participants do not just use an agreement token but employ some interactional resources to display mutual agreement so that the change can be overtly confirmed and a joint decision-making process can be followed. To begin with, SA does not just use the compliance token "EVET/*yes*" to show her agreement with the suggestion but extends her turn to offer a (re)formulation of the suggested change (Heritage & Watson, 1979). In doing this, she explicitly indicates the kind of change she is requested to make in the option (i.e., *prefer'ü atıyım/let me omit "prefer"*) while crossing out the verb "*prefer*" as suggested. In this way, she not only displays a stance of acceptance of the suggestion but also demonstrates and ensures correct understanding of the change so that it can be noted correctly on the exam paper. Secondly, DU and GE's modified repeats (Kim, 2002; Stivers, 2005) of a prior utterance (i.e., *prefer'ü at/omit "prefer"; yok yok öyle bi bilgi yok/no no there is no such information; yok/no*) also display and demonstrate their shared understanding of a claim made in the IR interaction and overtly confirm it.

Thirdly, DU and SA's supportive and complementary explanations in relation to SA's account in lines 20 and 24 not only show their agreement and affiliation with each other's stance but also bring in additional evidence that further confirms the change in the item. Upon SA's general account about the absence of the information in the new distractor in the reading text in line 13, DU provides a more specific explanation by examining the text in terms of the information in the new option and stating that the text includes no information about the expensiveness or cheapness of the products that are bought as mentioned in the option. DU's account is then extended and completed by SA, who this time checks to see what the text says about shopping and remarks that it refers to doing more shopping but not to the products' expensiveness. In this way, the participants collaboratively provide an account for the unlikelihood of the revised distractor as a correct answer and ensure shared understanding in the group.

Finally, DU and GE respond to SA's confirmation-seeking question not simply with a "yes" or a head nod but then use repetitional responses (Heritage & Raymond, 2012) to explicitly acknowledge that the distractor option becomes wrong (i.e., *yanlış oluyo/it becomes wrong*; *yanlış/wrong*). Like the other interactional resources, this also displays mutual agreement and ensures shared understanding with respect to the change, making the participants' stance clearer. By means of such interactional resources as modified repeats, additional accounts and repetitional responses, the participants also claim their epistemic right to confirm an assertion, indexing epistemic authority and control over the matter at hand since they all have equal epistemic access to the test material, and are in K+ position in terms of their English language expertise and since they share a common ground in terms of the institutional expectations with regard to test construction (Heritage & Raymond, 2005; Stivers, 2005).

Regarding the institutional role distribution of the participants in extract 4c, there are indications of tester and teacher roles and responsibilities in this IR interaction. To begin with the tester roles, it is seen that as the writer of the test item who is accountable for the test, the tester SA carefully evaluates the change in terms of not only the violated testing principle (i.e., more than one correct answer) but also other testing principles that are important for multiple-choice item construction (i.e., length of options) and monitors the item change reviewing process accordingly. In this respect, before finalizing the item, she challenges the new option in terms of its length and playing the devil's advocate, she seeks

the group's final confirmation about whether the information in the distractor option is certainly wrong, making sure that the revision of the item is complete and the item is in line with the testing principles. Furthermore, by offering a (re)formulation of the suggested change and thus, explicitly articulating the kind of change she is to make in the revised option, SA ensures correct and mutual understanding of the second change that the group decides on in phase RC. The tester's such an act of (re)formulating the prior turn of suggestion functions as a final check for the item and prevents an incorrect noting of the change on the test material, orienting to her role as the tester who is accountable for the test and who is responsible for producing the final product. In addition, SA's non-verbal behaviour of marking the change on her exam paper by crossing out the verb "*prefer*" from the new option as well as the other participants' instructions to the tester about what she should do exactly (i.e., to say "*buy*" (line 4), not to write "*prefer*" (line 6), to say "*customers buy expensive products*" (line 7), to omit "*prefer*" (line 9)) to shorten the option with a shift of gaze to her both orient to SA's responsibility and obligation to finalize the test item and make it ready for administration.

As for the teacher roles, the reviewer DU displays an orientation to her role as the teacher by evaluating the revised item from the students' perspectives (i.e., *bakacak incelicek/s/he will look and examine*). Using the third-person singular pronoun to refer to the students, she comments on the difficulty level of the item as the students will have to carefully examine the reading text to be able to answer the item. This implies the reviewers' epistemic right, responsibility and obligation (Heritage, 2013) to review the item with consideration of the abilities of the target group of students.

Despite an orientation to the institutional role distribution of the participants, it is seen that they share the responsibility for reviewing the change in the problematic item. They all review the change carefully by collaboratively providing a final round of accounts for it, offering a suggestion, evaluating the revised item in terms of different testing principles, and displaying and demonstrating mutual agreement and understanding. In addition, a joint reviewing process is encouraged by SA's confirmation-seeking question and her invitation to recheck the unlikelihood of the revised distractor (i.e., *yanlış oluyo di mi? hocam/it becomes wrong doesn't it miss?*) with a shift of gaze to the other participants. In doing this, she does not just wait for the other participants to do the

reviewing of the change as made relevant next by her question but she provides the first account before the other participants respond to her confirmation request. Thus, all the three participants actively and critically think about the revised item in and through talk, follow a joint decision-making process in finalizing the item, and eventually reach a shared understanding about the change. Their equal active engagement is also evident in their gaze behaviour as they continuously look at their exam papers, constantly reviewing the change and checking the appropriacy of the new distractor (Goodwin, 1981; Rossano, 2013). For instance, GE turns over the page in line 29 to look at the reading text upon the other participants' account with reference to it and DU shifts her gaze to SA's paper in lines 1 and 7 to check the length of the option before searching for a suggestion to shorten it. This is because they all have equal epistemic access to the available resources and thus, they assume equal epistemic state and authority.

Extract 10c presents the phase of RC in the reviewing interaction on the item problematized in extract 10a (i.e., phase PI) and revised in extract 10b (i.e., phase SC) to overcome the problem. It is to be remembered that in this extract, the teacher YU teaching at B1 plus level does the reviewing of a story exam item and provides his feedback to the tester ME. This item appears in a matching test where the students are expected to match the descriptions of four characters from the story they have read beforehand with one of the six character names given in a box. In the first segment of the extract (i.e., extract 10a, see Section 4.1), YU problematized the description of one particular character named Lizzie, claiming that Lizzie was one of the minor characters in the story whom he had not focused on while reading the story and thus, he was unable to match the statement with the expected answer "Lizzie". Addressing this problem, the IR group found an alternative character name from the story and formulated a description for it in the second segment of the extract (i.e., extract 10b, see Section 4.2) through a suggestion and decision making process. As a result, the problematic character name "Lizzie" was replaced with "Jeremy Stickles" and the description "*is the spy who helped John with his soldiers*" was formulated for the new character name in phase SC in extract 10b. However, the reviewing interaction for this particular item does not come to end when the item is revised but continues for further review of the accepted change this time in the following segment of the extract (i.e., extract 10c) to make sure that the change is complete and the problem is solved and that it raises no other problems that could go unnoticed in the item. This extract was also presented at a data session. The analysis provided here is, thus, supported by the analysis carried out at this data session.

**Extract 10c. Lizzie, Segment c**

- 1 YU: seçenek↑ler (0.5) kalabilir bu şekil[de  
**the options can remain like this**  
+ both look at their exam papers
- 2 ME: [aynen öyle.  
**that's right**
- 3 (1.2)
- 4 YU: >ama işte< burda şu anda şey oldu  
**but you see now here it turned out thing**  
+ points his pencil at the box of character names on his paper  
+ ME shifts gaze to YU's paper
- 5 (0.5)
- 6 er:[:
- 7 ME: [iki kız kardeş oldu  
**it turned out to have two sisters**  
+ shifts gaze back to her paper
- 8 YU: evet [bu ] kız kardeşleri biliyolar  
**yes they know these two sisters**
- 9 ME: [ ° mu ° ]  
**did it**
- 10 (1.3)
- 11 YU: jeremy yeri↑ne  
**in place of "jeremy"**  
+ turns to his copy of the story
- 12 1# (0.8)
- 13 ME: ya da şey olabilir (.) er  
**or it could be this thing**
- 14 (0.5)
- 15 şey lizzie (.) yerine  
**well in place of "lizzie"**
- 16 (0.4)
- 17 başka birinden bahsedebiliriz.  
**we can mention somebody else**  
+ shifts gaze to the copy of the story in YU's hand
- 18 (0.3)
- 19 başka (.) [(inaudible)  
**another**
- 20 YU: [bet mi vardı sonunda  
**was there "bet" in its end**
- 21 ((ME produces thinking face while YU continues to  
search through the story book for 2.5s))
- 22 betty mi vardı =bet-  
**was there "betty"**
- 23 (0.7)

24 [şeyin sonunda ]  
*in the end of the thing*

25 ME: [şeyi veri↑yim mi?]  
*shall I give this thing?*

26 (0.2)

27 er: burda (.) counsellor da verebilirim<sup>#1</sup>  
*here I can also give "counsellor"*  
 + turns gaze back to her paper  
 #1(lines 12-27) YU searches for a character name in his copy of the story by turning over the pages

28 ((YU shifts gaze to his exam paper, producing thinking face for 1.3s))

29 lizzie yerine  
*in place of "lizzie"*  
 + shifts gaze to YU, then back to her paper

30 (2.2)

31 YU: olabilir =hani erkek olması da↑ha  
*it is possible, well its being a man is more*

32 ME: huh-huh

33 YU: avantaj  
*advantageous*

34 ((ME crosses out the name "Lizzie" in the box on her paper))

35 çünkü  
*because*  
 + ME replaces "lizzie" with the distractor "counsellor" in the box of character names on her paper

36 (1.5)

37 erkek (0.2) soruyoruz ya üç tane  
*you know we ask for three men*

38 (0.3)

39 ME: co↑unsellor olsun o zaman  
*then let it be "counsellor"*  
 + turns gaze to YU

40 YU: huh-huh  
 + maintains thinking with his eye at the test material and his right hand holding the pages of the story he has already scanned through

41 (2.0)

42 ME: bi asker ismi daha vardı da şu an hatırlamıyorum ben  
 =böyle yine iki isimli  
*there was also a soldier name but I cannot remember now, again with two names like this*

44 ((ME produces thinking face for 4.0s while YU first produces thinking face and then looks at the pages of the story that he has come up to while scanning

through it))

45 judge er (.) refer (0.3) [son var  
*there is*  
+ shifts gaze to YU

46 YU: [evet judge  
*yes*  
+ shifts gaze to ME

47 evet (.) ama o (.) hani  
*yes but that is you know*  
+ turns gaze to his exam paper

48 ju[dge (.) hani ona judge (.) >diyebiliriz<]  
*you know we can refer to it as "judge"*  
+ flips the pages of the story that he has  
been holding in his right hand after scanning through them

49 ME: [judge (.) evet onu bilirler ]  
*yes they would know that*

50 (1.0)

51 o za[man °counsellor mantıklı°  
*then "counsellor" is plausible*  
+ shifts gaze to her paper

52 YU: [bence bu şekilde kalabilir  
*I think this could remain like this*  
+ shifts gaze to his exam paper

53 =[yapabilirler yani  
*I mean they can do it*

54 ME: [tamam  
*okay*

As YU and ME are both looking at their exam papers, extract 10c begins with YU's initiation of the item change reviewing process by orienting to the options (i.e., the extra character names given in the box) to check them with a rising intonation on the final syllable of the word "seçenekler/*options*" and then making a positive evaluation after a gap of brief silence. He comments that the options given in the box can stay as they are after the change made in one of the character names in the matching test. ME conveys an affiliative stance and displays a strong alignment with YU in line 2 via her use of the agreement token "aynen öyle/*that's right*". However, following a substantial lapse, YU recognizes a problem and attempts to challenge the options by launching a *but*-prefaced turn which projects the announcement of a problem in the revised item. His turn is embodied as he points his pencil at the box of character names on his paper. This draws ME's attention as she shifts her gaze to YU's paper. YU's challenging acts as a counter-claim to his own argument about the status of the options.



Just as YU tries to find a way to express the problem in the options, which is evident in the intra-turn pause and elongated hesitation marker in line 6, ME employs an anticipatory completion (Lerner, 1996a) in line 7 in the form of a candidate evaluation by shifting her gaze back to her paper and then requests YU's confirmation in line 9 with a soft tone of voice. This is met with a responsive uptake by YU, who displays agreement with ME's anticipatory completion and overtly confirms it through a compliance token (*evet/yes*) followed by a complementary account for the presented challenge in line 8. While ME points out that the options have two sister names, YU makes the problem with respect to the two sister names clearer by explaining that the students would know those names, orienting to their assumed knowledge about the characters in the story and thus, to the violation of the testing principle about the plausibility of the options as the students can easily guess and eliminate the sister names given as extras. In this way, the participants turn out to challenge the revised item collaboratively.

Following a gap of silence, turning to his copy of the story, YU initiates a suggestion-making sequence in line 11 to overcome the announced problem. He first locates the trouble source by indicating the character name that is to be replaced (i.e., *Jeremy*) and starts to search for an alternative character name in his copy of the story by turning over the pages. ME takes over in line 13 to proceed with the suggestion-making sequence initiated by YU as she launches a preface that projects the change process in the item (i.e., *ya da şey olabilir/or it could be this thing*). However, she then uses an embedded correction and performs an other-initiated repair action with respect to the problematic option by announcing the character name to be actually replaced with an emphasis on it, which shows that it is "*Lizzie*" and not "*Jeremy*" as indicated by YU. This is followed by her suggestion to replace "*Lizzie*" with a different name in line 17 as one possible way of solving the problem with an inclusive language (i.e., *bahsedebiliriz/we can mention*) by shifting her gaze to the copy of the story in YU's hand.

Failing to accurately recall the name of one of the characters that appears in the end of the story, YU carries out a preliminary inquiry (Pomerantz & Heritage, 2013) between lines 20 and 24 by orienting to ME's knowing status with regard to the story and asking her about this name. In line 21, while YU continues to search through the story book for alternative options for 2.5 seconds and tries to recall the name of this particular character (i.e., *bet*,

betty), ME produces thinking face. Without orienting to YU's preliminary inquiry, coming up with an alternative, ME applies a preface in line 25 with a rising intonation by using the first person singular pronoun (i.e., *şeyi veri↑yim mi?/shall I give this thing?*), which projects an upcoming suggestion. Turning her gaze back to her paper and using an elongated hesitation marker that demonstrates her online thinking process about her suggestion, she then suggests the character "*counsellor*" as an alternative option in line 27 by using the first person singular pronoun (i.e., *verebilirim/I can give*) and once again addresses the trouble source that it is to replace in line 29 (i.e., *Lizzie*) with a shift of gaze to YU and then back to her paper. She also downgrades her suggestion by presenting it just as one possibility for the other participant to consider (i.e., *verebilirim/I can give*) and requesting YU's confirmation in the preface (i.e., *veri↑yim mi?/shall I give*).

Upon ME's announcement of the suggestion, YU shifts his gaze to his exam paper, producing thinking face for 1.3 seconds. Following a lapse of 2.2 seconds after ME's turn where she reminds the trouble source "*Lizzie*" in the options, YU shows an acceptance of the suggestion as one possibility in line 31 (i.e., *olabilir/it is possible*); thereby not fully agreeing with it and deciding on it but projecting the continuation of his search for different alternatives. YU then rushes to provide a candidate explanation for ME's suggestion between lines 31 and 37 with an inclusive language (i.e., *soruyoruz/we ask*) by once again orienting to the testing principle regarding the plausibility of the options and this time indicating the assurance of this principle since three of the character descriptions ask for a man name and the suggested name is a man, which is more advantageous in terms of making the options plausible. As YU formulates his account for the new alternative, ME displays agreement with him through the acknowledgement token "huh-huh" and then crosses out the name "*Lizzie*" in the box of character names on her paper, replacing it with the distractor "*counsellor*".

Upon YU's completion of his account for the new alternative, turning her gaze to YU, ME repeats the suggestion to receive YU's final confirmation for it and to make the final decision with a rising intonation on the second syllable of the alternative "*co↑unsellor*". In line 40, YU displays agreement via the acknowledgement token "huh-huh"; however, it

is not a strong agreement as he maintains thinking with his eye at the test material and his right hand holding the pages of the story he has already scanned through, which signals his continued consideration of the matter at hand. As a result, after a substantial lapse, ME comes up with another alternative (i.e., a soldier name) but she claims failure to accurately recall it and then produces thinking face for 4.0 seconds while YU also first produces thinking face and then looks at the pages of the story that he has come up to while scanning through it.

As they think of different alternatives, ME recalls a different man name from the story in line 45 (i.e., *judge referson*) and presents it as another alternative to consider. Just prior to the completion of her turn, she maintains mutual gaze with YU, who acknowledges the existence of such a character in the story through the compliance token “*evet/yes*” in line 46 but then displays dispreference and disagreement with this alternative as projected by his use of “*ama/but*” with a shift of gaze to his exam paper. YU challenges this name by referring to its implausibility using an inclusive language (i.e., *diyebiliriz/we can refer*) since the term “*judge*” makes it too obvious. Through the end of his unit of talk, he starts to flip the pages of the story that he has been holding in his right hand after scanning through them, which further indicates his check of all the possible alternatives in the story. In overlap with YU’s turn, ME employs an anticipatory completion (Lerner, 1996a) in line 49 and completes YU’s ongoing TCU of challenging the suggestion, which acts as a supportive and complementary explanation for the inappropriacy of the offered suggestion “*judge referson*” as a distractor option and shows her agreement with YU. She explains the implausibility of the presented alternative (i.e., *judge referson*) by orienting to the students’ assumed knowledge of the character and thus, indicating their quick elimination of this option.

After a gap of silence, drawing a conclusion based on YU’s account, the decision on the change is made clearer as ME reannounces the first offered suggestion “*counsellor*” to receive YU’s confirmation and comments on its plausibility with a soft tone of voice by shifting her gaze to her paper. This time she makes a stronger evaluative stance towards her suggestion and displays certainty. In line 52, overlapping ME’s turn-at-talk, YU also makes the final comment and confirms the change in the item. Shifting his gaze to his exam paper, YU then rushes to provide an account for the revised item from the students’ perspective by

explaining that they can now do the matching of the new item and hence, orienting to the students' assumed skill. In this way, he makes a reference to the main issue in the item (i.e., the character "*Lizzie*" being a minor detail in the story) and thus, the testing principle that each item should be based on important content. Consequently, he reveals the criterion for the revision of the item in and through the talk-in-interaction and concludes that the problem identified in the phase of PI is solved. This is received with ME's agreement through her use of the minimal expansion acceptance token "*tamam/okay*", which also proposes sequence closure. As a result, the revision of the item is finalized and the reviewing interaction for this particular item ends.

Extract 10c demonstrates how the participants review the item that is revised in the phase of SC to replace the minor character "*Lizzie*" in one of the matching items with a more important content from the story and how they finalize the reviewing of this item. The item change reviewing process begins with the participants' orientation to the extra character names given in the box in an attempt to make sure that the change fulfils not only the violated testing principle but also other testing principles, causing no additional problems and their recognition of a problem in the distractor options as a result of the replacement of the problematic character "*Lizzie*" with "*Jeremy*" in phase SC. The participants realize that the extra character names end up having two sister names which the students already know and can, thus, easily guess and eliminate (i.e., *iki kız kardeş oldu/it turned out to have two sisters; bu kız kardeşleri biliyorlar/they know these two sisters*). Orienting to the students' assumed knowledge about the characters in the story, they find out that the testing principle that "options should be reasonably attractive and plausible" (Carr, 2011; Haladyna & Rodriguez, 2013; Hatipoğlu, 2009; Heaton, 1990) is violated. Based on this, the matter at hand is displayed and regarded as assumed knowledge with respect to the testing principles adopted within the institutional context in question since the problem is defined collaboratively through ME's anticipatory completion in the form of a candidate evaluation and YU's complementary explanation. It is also not further questioned and discussed as the group directly orients to a search for a change in the options to overcome the problem, which signals their already acceptance of the presented challenge.

The interactional action that plays an important role in the item change reviewing process in this extract is, thus, the participants' collaborative challenging of the options. The act of challenging leads to the recognition of a problem that emerges in the matching test after the problematic item undergoes a change, shifting the focus of the interaction from the testing principle concerning the content that is tested in this item onto the testing principle about the plausibility of the alternative character names given in the box. Consequently, the act of challenging determines the rest of the interaction in phase RC in extract 10c and one of the character names appearing as extra in the box (i.e., *Lizzie*) after the major change made in the test is eventually substituted with a man name from the story to prevent other emerging problems and to comply with the testing principles. This indicates that language test IR interactions require continuous reviewing and editing since a major change (i.e., replacing one of the matching items with another character name and description) might lead to other changes (i.e., changing one of the extra character names in the box) to make the revision of the item complete. As a result, the findings show that the change process in the revision of a problematic item in an IR interaction can be much more complex and dynamic, requiring a careful and critical eye.

The analysis has shown that a new suggestion-making sequence takes place in the phase of RC following the same structural organization that appears in the phase of SC. In a similar vein, the interactional actions of applying a preface, offering two suggestions subsequently, challenging a suggestion, downgrading a suggestion, and giving account for a suggestion that were revealed in the phase of SC also appear in this extract. In this extract, four different boy names are suggested from the story to replace the extra character name "*Lizzie*" (i.e., bet/betty, counsellor, a soldier name, judge referson) and the group decides on the second one (i.e., counsellor). As in the phase of SC, at first two suggestions are made subsequently whereby YU and ME provide their suggestions respectively without orienting to the one that has already been made (i.e., bet/betty). This could be either because the two participants both seek alternatives and share with the group the name that they can recall from the story so that they can make a decision accordingly or because ME is after better alternatives, proceeding with her search for an alternative, which can be interpreted as a display of dispreference with the prior suggestion offered by YU. As a result, the participants orient to ME's suggestion (i.e., counsellor) and make a positive evaluation of it. What seems to play an important role in the acceptance of ME's suggestion is YU's account-giving in the form of a candidate explanation, which makes it clear that different from the problematic option

“Lizzie”, “counsellor” is a man name and thus, is more plausible since three of the character descriptions in the test ask for a man name. In this respect, “counsellor” seems to prove the validity of the testing principle violated by “Lizzie”.

Despite the group’s positive stance towards the alternative “counsellor”, the suggestion-making sequence is still not steered to the end point and the final decision since YU does not show a strong agreement with ME’s suggestion and his “huh-huh” falls short. He regards it just as one possibility to consider and maintains thinking with his eye at the test material and his right hand holding the pages of the story he has already scanned through, which signals his continued consideration of the matter at hand and his check of all possible alternatives before making the final decision. This leads ME to come up with two more alternatives (i.e., a soldier name, judge referson). As in the case of the first suggestion offered by YU, no orientation is shown to the soldier name as ME cannot recall the exact name of the soldier and displays uncertainty about it. As for the final suggestion (i.e., judge referson), it is collaboratively challenged and rejected with reference to the students’ assumed knowledge of this character and thus, its failure in complying with the testing principle about the plausibility of the option.

Such a process of suggestion-making shapes the talk-in-interaction in the phase of RC to a considerable extent and demonstrates how the participants interactionally deal with the problem identified in the options after the revision of one of the matching items. It is found that multiple alternatives are considered before deciding on a change in the problematic option “Lizzie” as the participants are constantly in search of a better alternative. Their articulation of every new idea that comes to mind and its discussion in and through talk is important in deciding on a mutually acceptable change, enabling the group to achieve consensus on the change and to eventually confirm it “in open discussion, through a process of dialogue and disagreement” (Fulcher & Davidson, 2007, p.21), which makes an important contribution to the validation process. It is only after they consider all alternatives from the story that they find out “counsellor” is the best possible character name to include as an extra in the options, which was already proved by YU’s account-giving. Therefore, as it was found in the phase of SC, the suggestion that is justified based on evidence is more likely to be accepted.

Regarding the act of suggesting a change, as in the phase of SC, ME downgrades her suggestion by presenting it just as one possibility for the other participant to consider (i.e., verebilirim/*I can give*; bahsedebiliriz/*we can mention*; şey olabilir/*it could be this thing*) and requesting YU's confirmation in the preface (i.e., veri↑yim mi?/*shall I give*). In this way, ME does not implicate epistemic priority as the first provider of the suggestion (Heritage & Raymond, 2005) but instead invites collaborative thought on the issue and solicits mutual agreement, sharing the responsibility for the decision. Thus, ME might not want to block the other participant's contributions by being too assertive about her suggestion but rather pave the way for other alternatives and for a careful thought and decision-making on the extra character names. This is because the participants have equal epistemic access to the referent situation (e.g., the test material and the story) and share a common ground in terms of the testing principles for language test construction. On the other hand, ME's subsequent display of certainty with her suggestion through a stronger evaluative stance after eliminating other possible alternatives (i.e., co↑unsellor olsun o zaman/*then let it be "counsellor"*; o zaman °counsellor mantıklı°/*then "counsellor" is plausible*) leads to the acceptance of her suggestion by the group. Furthermore, as a commonly used interactional resource in the phase of SC, ME applies a preface twice in the talk-in-interaction (i.e., ya da şey olabilir/*or it could be this thing*; şeyi veri↑yim mi?/*shall I give this thing?*), which projects the announcement of her upcoming suggestion.

The participants use some interactional resources to display mutual agreement through a collaborative process and to ensure shared understanding about the change in the item so that a joint decision-making process can be followed. To begin with, ME does it through her anticipatory completions (Lerner, 1996a) in lines 7 and 49. Her first anticipatory completion (i.e., iki kız kardeş oldu/*it turned out to have two sisters*) includes a candidate evaluation of the options where she points out her analysis that the extras turned out to have two sister names. In her second anticipatory completion (i.e., judge (.) evet onu bilirler/*yes they would know that*), she completes YU's ongoing TCU of challenging one of her suggestions (i.e., judge referson), which acts as a supportive and complementary explanation for its implausibility as a distractor option. In both cases, the participants end up employing the act of challenging

through collaborative turn sequences, discussing the addressed issues jointly and thus, they demonstrate their shared understanding about the matter at hand, approaching it as assumed knowledge with respect to the testing principles adopted within the institutional context in question.

In addition, YU displays agreement with ME through his supportive and complementary explanation in line 8 as well as his candidate explanation for ME's suggestion between lines 31 and 37. In the case of the former, YU does not just use a compliance token to display agreement with ME's anticipatory completion but goes beyond it to provide an additional explanation for the issue raised by ME with regard to the two sister names (i.e., *bu kız kardeşleri biliyolar/they know these two sisters*), which not only makes the problem in the options clearer but also supports ME's candidate evaluation. As a result, they turn out to challenge the revised test collaboratively. As for YU's candidate explanation for ME's suggestion "counsellor", it shows his stance of acceptance of the suggested change and establishes shared understanding about it by presenting the underlying reason for the change (i.e., *erkek soruyoruz ya üç tane/you know we ask for three men*) and indicating why ME's suggestion is a more plausible and valid option (i.e., *erkek olması daha avantaj/its being a man is more advantageous*). By means of such interactional resources, the participants do not only show mutual agreement but also claim their epistemic right to confirm an assertion, indexing epistemic authority and control over the matter at hand since they all have equal epistemic access to the story and are in K+ position in terms of their testing knowledge (Heritage & Raymond, 2005; Stivers, 2005).

Regarding the institutional role distribution of the participants in extract 10c, there are indications of tester and reviewer roles and responsibilities in this IR interaction. To begin with the tester roles, it is seen that different from the reviewer YU who searches for alternative character names through the copy of the story in his hand by flipping the pages, the tester ME recalls the names from memory and is the one who often comes up with alternatives before YU does so. This orients to her right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the story as the writer and thus, the 'owner' of the item who has analyzed the story in detail. This is the reason why YU asks ME about the exact name of one of the characters (i.e., *bet mi vardı sonunda/was*



*there "bet" in its end; betty mi vardı =bet-/was there "betty")* while carrying out a preliminary inquiry into this character before offering it as a suggestion, which ascribes an epistemic authority to ME. In a similar vein, the analysis has revealed that the tester usually takes over and completes YU's turns following his pauses and elongations in, for example, lines 6, 12 and 23. This further indicates the tester's greater command of the story and her control over the test material.

The tester ME also monitors the change process by reminding the trouble source "Lizzie" in line 29 as well as her suggestion "counsellor", which was evaluated positively, after the other possible alternatives are eliminated (i.e., *coşuncsellor olsun o zaman/then let it be "counsellor"; o zaman °counsellor mantıklı°/then "counsellor" is plausible*) and directing the trajectory of the interaction towards the final decision. By bringing her suggestion to the forefront of the group's attention with a shift of gaze to YU, for instance, in line 39, ME also attempts to receive YU's confirmation with regard to the decided change after considering multiple alternatives. In those ways, ME ensures mutual understanding and agreement in the group before finalizing the test and ending the IR interaction, which orients to her role as the tester who is accountable for the test and who is responsible for producing the final product. Additionally, ME's non-verbal behaviour of noting down the change on the test material in line 35 as well as her use of the first person singular pronoun while offering her suggestion as an extra name to give in the options in lines 25 and 27 orients to her responsibility and obligation to finalize the test and make it ready for administration.

As for the reviewer roles, YU orients to his role as the problematizer of the item by making a reference to the main issue he raised in phase PI and thus, the testing principle that each item should be based on important content (i.e., *yapabilirler yani/I mean they can do it*). He claims that the students can now do the matching of the new item, which at the same reveals the criterion for the revision of the item through the talk-in-interaction and concludes that the problem identified in phase PI is solved. This signals his responsibility to remind the IR group of the main issue and to establish common ground on the underlying reason for the replacement of the problematic item at the point when the participants shift their focus to the options of the revised test while reviewing it. Hence, as the problematizer of the item, he makes sure that the problem is addressed and is clearly understood by the

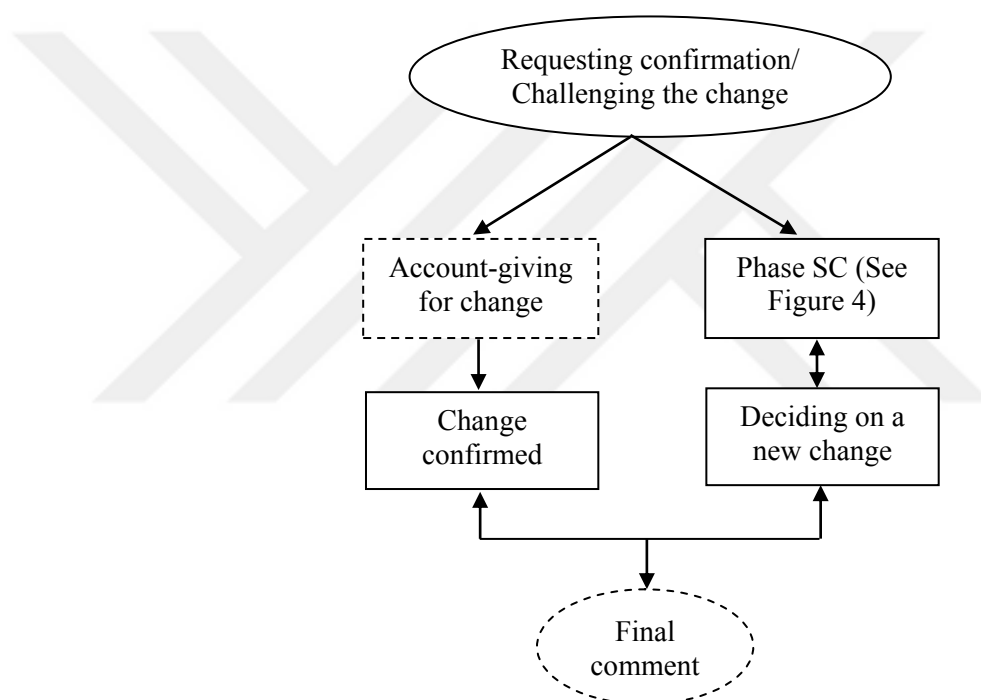
group. Moreover, YU tries to recall names from the story as in lines 20 and 22 and displays knowledge of the presented names by, for instance, providing account for “counsellor” and challenging “judge referson”. This demonstrates that he has fulfilled his responsibility to read the story before coming to the IR session for the story exam, orienting to his K+ status about the issue (Heritage, 2013).

Finally, YU makes his role as a teacher explicit by referring to the students’ assumed knowledge and skills (i.e., *biliyolar/they know; yapabilirler/they can do it*). Using the third-person plural pronoun to refer to the students, he claims their knowledge of the extra character names in the box in line 8, presenting their knowing as negative to refer to the implausibility of the options. In line 53, he once again uses the third-person plural pronoun to refer to the students and this time comments on their assumed ability to do the matching of the revised item, which now tests the students’ comprehension of an important content from the story unlike the minor character “Lizzie” originally tested. This implies the reviewers’ epistemic right, responsibility and obligation (Heritage, 2013) to review the item with consideration of the abilities and knowledge of the target group of students.

Despite an orientation to the institutional role distribution of the participants, it is seen that they share the responsibility for reviewing the test after the problematized item is revised. The two participants both orient to the alternative names given in the box and review them carefully by collaboratively challenging them, offering suggestions and evaluating them, and displaying and demonstrating mutual agreement and understanding. In addition, a joint decision-making process is encouraged by ME’s downgrading of her suggestion and her reannouncement of it with a shift of gaze to YU to elicit his confirmation after a consideration of different alternatives. In addition, ME does not just wait for YU to offer alternatives or make an evaluation but takes over to complete YU’s turns or to proceed with the sequences initiated by YU, following a collaborative process. The participants’ equal active engagement is also evident in their gaze behaviour as they continuously look at their exam papers only sometimes directing their gazes at each other for a short time (Goodwin, 1981; Rossano, 2013). YU, for instance, constantly searches through the story book as the primary source by turning over the pages and the participants both produce thinking face throughout the talk-in-interaction as they think about the character names in the story. This is because the participants assume equal epistemic state and authority since they have equal

epistemic access to the available resources (e.g., the story, test material) and share a common ground in terms of the testing principles for language test construction. Finally, the participants' use of an inclusive language during the suggestion-making sequence (i.e., *bahsedebiliriz/we can mention*; *soruyoruz/we ask*) represents their shared responsibility for the change process in the item and shows that the revised test is the product of the whole group.

#### 4.3.1. Summary of Interactional Patterns in Phase RC



\*See Appendix C for the meaning of flowchart symbols

**Figure 5. Overall Structural Organization of Phase RC**

Regarding the findings for the overall structural organization of the phase of RC in the language test IR interactions analysed for this study, it is seen that this phase is about the participants' review of the change that is decided in the phase of SC to overcome the problem identified in the item in the phase of PI. In this phase, the IR group checks and finds out whether the revised item complies with the testing principles for language test

construction, addressing the violated testing principle and causing no additional problems that could go unnoticed with the revision of the item. Such a stage is vital considering the significance and the effect of the product that is produced in this task-oriented IR interaction. This phase is found to be operationalized through a particular recurring interactional pattern as illustrated in Figure 5 and this pattern is often repeated several times in the extracts (i.e., 1c, 3c, 4c) as the participants start over the item change reviewing process to do a final check from different perspectives. For instance, in extract 4c, once the revised item is challenged and a new change is made to solve the emerging problem, another reviewing sequence is initiated to check the revised item this time in terms of the main issue raised in the phase of PI (i.e., the violated “only one correct answer” principle).

Considering the overall structural organization of phase RC, the sequence-initiating actions in this phase involve the participants’ confirmation requests or their challenging of the change, which leads the group to reconsider the revised item. To begin with the former, it is seen that through their confirmation requests which often appear together with a tag question (extracts 1c, 3c, 4c), the participants mobilize support for their claim regarding the change made in the item (Heritage, 2012). They make sure that the change fulfils the testing principles, causing no additional problems and establish shared understanding with regard to it before finalizing the item and ending the reviewing interaction for this particular item. The participants’ confirmation requests for the change enable them to check several points: (1) the group’s preference for the change (i.e., *↑böyle mi daha iyi? designed <by'lı mı> kalsın/is it better like this? should it stay with “designed by”, extract 1c*), (2) their correct understanding of the change (i.e., *meat (.) dedik di mi? =milk değil?/we said “meat” right? not “milk”, extract 1c*), and (3) the incorporation of the testing principle that is claimed to be violated in the first phase of the IR interaction (e.g., more than one correct answer) (e.g., *yanlış oluyo di mi? hocam/it becomes wrong doesn't it miss?, extract 4c*) as well as other testing principles for language test construction (e.g., *yoktu di mi? °hiç o°/it was not present at all was it?*). Those confirmation requests invite the group to re-evaluate the change and hence, lead to a collaborative process of reviewing the change, initiating a discussion about its appropriacy.

Another sequence-initiating action that plays an important role in shaping the talk-in-interaction in the item change reviewing process is challenging the change as seen in extracts 1c, 2c, 4c, and 10c. While reviewing the change in the test item, the participants recognize a problem and challenge the revised item with reference to the testing principles for item writing. It is seen that in the extracts where the change is reviewed and challenged, the participants identify a new problem that emerges in the item as a result of its revision in phase SC and that violates a testing principle that is different from the one raised in phase PI. As Table 20 illustrates, in extract 1c, since the item in question asks for an advantage *not* mentioned in the text, DE challenges the idea of “*meat*” in the newly formulated correct option in terms of its relevance to healthy food mentioned in the reading text (i.e., *ama meat de o zaman-/but then “meat” too-*). This addresses the violation of the avoidance of “no answer” principle rather than the main problem with respect to the plausibility of the correct option, which has already been solved in phase SC. Similarly, in extract 2c, as a result of the replacement of the problematized verb “*cover*” with “*change*” to be transformed into its past tense form in the tense passage, the participants realize that the new verb “*change*” does not fit grammatically well in the sentence (i.e., *change sanki (.) >anlam olarak da< oturur gibi di mi?/it is like “change” seems to fit in terms of meaning too, doesn’t it?*). This shows an orientation to a problem that emerges in the item after the verb of the target sentence undergoes a change, shifting the focus of the interaction from “test what you teach” principle as the main concern onto the grammaticality of the language used in the item.

As for extract 4c, through the act of challenging the change (i.e., *bu: >muhtemelen biraz< uzun olacak >diğerlerine göre< ama/but this will probably be a little longer compared to the others*), an orientation is displayed to the violation of a testing principle which states that “options should be similar in structure” (Fulcher, 2010, p.172) and more specifically that the length of options should be kept about equal (Haladyna & Rodriguez, 2013). As a result, the focus of the interaction is shifted from the content of the option and “only one correct answer” principle as the main concern onto its format. Finally, in extract 10c, as a result of the replacement of the minor character “*Lizzie*” with a more important character “*Jeremy*” from the story in one of the matching items in phase SC, the participants realize that the extra character names given in the box end up having two sister names which the students already know and can, thus,

easily guess and eliminate (i.e., *iki kız kardeş oldu/it turned out to have two sisters*; *bu kız kardeşleri biliyorlar/they know these two sisters*). Hence, orienting to the students' assumed knowledge about the characters in the story, they find out that the testing principle stating that "options should be reasonably attractive and plausible" (Carr, 2011; Haladyna & Rodriguez, 2013; Hatipoğlu, 2009; Heaton, 1990) is violated, which shifts the focus of the interaction from the testing principle concerning the importance of the content that is tested in this item onto the testing principle about the plausibility of the alternative character names given in the box as distractors.

**Table 20. Summary of Challenges Presented to the Change and Their Outcomes**

| Extract No | Decided Change in Phase SC                                            | Presented Challenge                                                                                                                                                                              | Outcome                                 |
|------------|-----------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------|
| 1c         | <i>"following a special diet consisting of meat"</i>                  | The possibility of being related to the idea of healthy food mentioned in the text as the item asks for information <i>not</i> mentioned in the text<br><br>(Avoidance of "no answer" principle) | Challenge rejected/<br>Change confirmed |
| 2c         | <i>"change"</i>                                                       | Not collocating with the preposition <i>"with"</i> in the target sentence, making it grammatically inaccurate<br><br>(Grammaticality principle)                                                  | Challenge accepted/<br>Problem solved   |
| 4c         | <i>"customers prefer to buy expensive products"</i>                   | Being longer than the other options<br><br>(The odd man out principle)                                                                                                                           | Challenge accepted/<br>Problem solved   |
| 10c        | <i>"Jeremy Stickles is the spy who helped John with his soldiers"</i> | The extra character names in the box turning out to have two sister names which the students already know and can, thus, easily guess and eliminate<br><br>(Plausibility principle)              | Challenge accepted/<br>Problem solved   |

Following the participants' confirmation requests or act of challenging the change as the sequence-initiating actions, the talk-in-interaction in RC is sequentially unfolded in two ways; first when the reviewed change is confirmed by the group (extracts 1c, 3c, 4c) and

second when the item undergoes another change through a new suggestion and decision-making process as a result of being challenged and disconfirmed (extracts 2c, 4c, 10c). In the first case, the sequence-initiating actions such as confirmation requests entail a final check of the revised item and allow the group to confirm the change often following a process of account-giving in support of it as in extracts 1c, 3c, and 4c. In addition, as seen in extract 4c, even though the participants' display of agreement with regard to the revised item does not project an expansion in phase RC, the act of account-giving for the already decided and noted change in the item might continue in a non-minimal post-expansion sequence (Stivers, 2013) through a collaborative process after the confirmation of the change (illustrated by the double-sided arrow in Figure 5).

The findings show the participants provide account for the change in the item collaboratively from different perspectives and confirm the revised item based on evidence. In other words, through account-giving, they make sure that the problem is solved and the testing principle in question is put to use. In extracts 3c and 4c, for instance, account is provided to show that the change ensures the violated “only one correct answer” principle (Heaton, 1990; Hatipoğlu, 2009). Furthermore, while giving account, the participants bring their professional expertise to the interaction and show their subject knowledge as language teachers (extract 3c) or present concrete evidence from the shared sources such as the reading text (extract 4c). In extract 3c, for instance, to prove that the new distractor option “*whereas*” is not possible as a correct answer, the change is evaluated from different dimensions based on the participants' language expertise; (1) the content and meaning of the target sentence (i.e., *zıtlık yok/there is no contrast*), (2) the punctuation rules for “*whereas*” (i.e., *noktadan sonra da olmaz whereas/“whereas” also wouldn't be okay after full stop*), and (3) the type of conjunction that is required in the target sentence (i.e., *ara bağlaç olması lazım/it should be a coordinating conjunction*).

Likewise, in extract 4c, the participants evaluate the appropriateness of the new distractor (i.e., *customers prefer to buy more expensive products*) in terms of its unlikelihood as a correct answer by presenting evidence from the reading text and providing account from general to more specific as they carefully re-examine the reading text in terms of; (1) the absence of the new point stated in the revised distractor (i.e., *çünkü böyle bi bilgi*

yok genelde/*because there is no such information in general*), and (2) what is stated in the text about the issue of the expensiveness or cheapness of the products people buy as stated in the revised distractor (i.e., *dařha çok alışveriş yapsın diyo ama expensive olması söz konusu deęil/it says s/he should do more shopping but its expensiveness is not the point in question*). Consequently, such a collaborative and comprehensive process of account-giving leads to the acknowledgement and approval of the change in open discussion as well as the establishment of shared understanding and mutual agreement on the revised item in the IR group.

On the other hand, in the second case when the change is challenged and is not confirmed by the group, a new suggestion-making sequence takes place following the same structural organization that has appeared in phase SC (extracts 2c, 4c, 10c). As a result, the group decides on a new change to overcome the problem that has emerged after the revision of the item and the item is eventually finalized. As Table 20 further shows, as a result of the act of challenging the change in the revised item, except in extract 1c where the presented challenge is rejected and the change is confirmed through account-giving as in the first case, the participants accept the presented challenge and move towards solving the emerging problem through a suggestion and decision making process in the rest of the interaction in phase RC. In fact, the matter at hand is displayed and regarded as assumed knowledge with respect to the testing principles adopted for language test construction within the institutional context in question since it is not further questioned and discussed as the group directly orients to a search for a change to overcome the problem, which signals their already acceptance of the presented challenge. In this respect, the participants' act of challenging the change determines the rest of the interaction in phase RC and generally another change process takes place to prevent other problems that would arise as a result of the major change made in the item and thus, to comply with the testing principles. This indicates that language test IR interactions require continuous reviewing and editing since a major change might lead to other changes to make the revision of the item complete. This could happen either within the phase of SC before the decided change is fully acknowledged and noted on the exam paper as seen in extracts 8b and 10b or in a separate phase as in the extracts in question (i.e., 1c, 2c, 4c, 10c) where the participants review the already noted change in terms of different testing principles before proceeding with the other items in the



test. Consequently, the findings show that the change process in the revision of a problematized item in an IR interaction can be much more complex and dynamic, requiring a careful and critical eye.

With respect to the new suggestion-making sequence that takes place in the phase of RC (extracts 1c, 2c, 4c, 10c), the interactional actions of applying a preface, offering two suggestions subsequently, challenging a suggestion, downgrading a suggestion, and giving account for a suggestion that were revealed in phase SC also appear in phase RC. In response to a challenge presented towards the change in the item, the participants initiate a suggestion-making sequence for another change process in the item to solve the emerging problem. An exception to this is extract 1c where a suggestion-making sequence is launched not in response to a challenge to overcome a problem but just to consider another alternative that has emerged in the talk-in-interaction for the already revised distractor option (i.e., *dairy products*). However, this suggestion is then challenged and withdrawn and thus, different from the other extracts, the act of offering a suggestion does not result in a new change in the item. Consequently, the suggestion-making sequence in extract 1c seems to appear as a subsidiary sequence within the item change reviewing process. This shows that the participants are constantly in search of a better alternative even if they have decided on a change in the problematized item.

On the other hand, as Table 21 illustrates, in extracts 2c, 4c, and 10c, suggestions are offered and a decision is made on a change to solve the problem that is identified in the newly revised item. While slight modifications are suggested for the changes already made in the problematized part of the items in extracts 2c and 4c (i.e., replacing the preposition in the target sentence, omitting one of the verbs in the distractor option), it is a different part of the matching test that undergoes a change in phase RC in extract 10c (i.e., one of the extra character names given in the box) as a result of replacing one of the matching items with another character name and description. In addition, in extracts 2c and 4c, the participants focus on one particular suggestion and quickly put it to use. However, in extract 10c, the suggestion-making process lasts longer as four suggestions are offered in total to replace the extra character name “Lizzie” in the box and the group decides on the second one (i.e., counsellor). As in the phase of SC, at first two suggestions are made subsequently without an orientation to the one that has already been made (i.e., bet/betty). This could be either because the two participants both seek alternatives and share with the group the name that

they can recall from the story so that they can make a decision accordingly or because they are after better alternatives, proceeding with their search for an alternative, which can be interpreted as a display of dispreference with the prior suggestion. In relation to this, despite the group's positive stance towards the second alternative (i.e., counsellor), the suggestion-making sequence is still not steered to the end point as the participants attempt to check all possible alternatives before making the final decision. It is only after they consider all alternatives from the story that they find out "*counsellor*" is the best possible character name to include as an extra in the options. Such a process of suggestion-making shapes the talk-in-interaction in phase RC to a considerable extent and demonstrates how the participants interactionally deal with the problem identified in the options after the revision of one of the matching items.

**Table 21. Summary of Suggestions and Decided Changes for the Problems in the Revised Language Test Items**

| Extract No | Problem in the Revised Item                                                                                                                         | Suggestions                                                                                         | Decided Change                              |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|---------------------------------------------|
| 2c         | The new verb " <i>change</i> " does not collocate with the preposition " <i>with</i> " in the target sentence.                                      | Replacing " <i>with</i> " in the original target sentence by " <i>using</i> "                       | " <i>using</i> "                            |
| 4c         | The newly formulated distractor option is longer than the other options.                                                                            | Omitting the verb " <i>prefer</i> " from the new option                                             | " <i>customers buy expensive products</i> " |
| 10c        | The extra character names in the box turned out to have two sister names which the students already know and can, thus, easily guess and eliminate. | 1. <i>bet/betty</i><br>2. <i>counsellor</i><br>3. <i>a soldier name</i><br>4. <i>judge referson</i> | " <i>counsellor</i> "                       |

What seems to play an important role in the acceptance of a suggestion is the act of account-giving. In extract 2c, CI's account based on her language expertise makes it clear that "*changed*" does not collocate with the preposition "*with*", causing the verb expression "*changed with it*" to sound weird (i.e., *changed with it bi garip oluyo/sounds weird*). It is only after her account following her suggestion that the other co-participants show an explicit agreement and display of understanding. Likewise, in extract 10c, YU's account-giving in the form of a candidate explanation with reference to a

testing principle (i.e., plausibility of the options) shows that different from the problematized option “Lizzie”, “counsellor” is a man name and thus, is a more plausible and valid option since three of the character descriptions in the matching test ask for a man name. Therefore, as it was found in the phase of SC, the suggestion that is justified based on evidence is more likely to be accepted.

The act of challenging an offered suggestion also takes place in phase RC during the suggestion and decision making process in the extracts analysed for this study. In doing this, the participants make a negative evaluation of the offered suggestion with reference to the testing principles for item writing as well as the students’ present state of L2 knowledge. In extract 1c, for instance, the participants challenge the new alternative that arises in and through the talk-in-interaction for the correct option (i.e., *dairy products*). DE questions the word “*dairy*” in terms of the students’ knowledge (i.e., *dairy’i bilemez↑se/what if s/he does not know “dairy”*) and BA challenges the idea of dairy products as she thinks that it would be associated with the idea of healthy food, being closely related to the information in the text and thus, ending up as an incorrect answer since the item asks for an advantage *not* mentioned in the text (i.e., *o ↑da (0.2) sağlıklı bi yiyeceğe: (.) götürebilir (.) fikren/it then might take us to the idea of healthy food*). While DE orients to the testing principle that states “the item should represent a specific content area covered in the course” (i.e., “test what you teach”) (Carr, 2011; Haladyna & Rodriguez, 2013), BA refers to the avoidance of “no answer” principle as a result of the revisions in the item (Hatipoğlu, 2009) despite the fulfilment of the plausibility principle as the main concern.

Likewise, in extract 10c, the participants collaboratively challenge the final suggestion (i.e., judge referson) offered for the problematized character “Lizzie” given as an extra in the box with reference to the students’ assumed knowledge of this character and thus, its failure in complying with the plausibility principle (i.e., *ama o (.) hani judge (.) hani ona judge (.) >diyebiliriz</but that is you know “judge” you know we can refer to it as “judge”*; *judge (.) evet onu bilirler/yes they would know that*). As a result of the act of challenging a suggestion, as it was found in phase SC, the participants (1) self-cancel and withdraw their own suggestion (extract 1c), (2) reject an offered suggestion, orienting to the other possible

alternatives (extract 10c), or (3) counter-challenge a claim made against a suggestion, rejecting and disproving the challenge presented towards an offered suggestion as in extract 1c (i.e., *aslında bilmesi lazım tabi çok (.) öğrettik ama/actually s/he should know of course we have taught it a lot but*) although the counter-challenge in this extract does not result in the acceptance of the suggestion due to the other problems that it would pose. The act of challenging a suggestion shows that the articulation of every new idea that comes to mind and its discussion in and through talk is important in deciding on a mutually acceptable change, enabling the group to achieve consensus on the change and to eventually confirm it “in open discussion, through a process of dialogue and disagreement” (Fulcher & Davidson, 2007, p.21).

With respect to the interactional act of suggesting a change for the problem in the revised item, it is further found that the participants produce their suggestions by either downgrading them (extracts 2c, 10c) through, for example, hedges (e.g., *belki/maybe*, extract 2c), confirmation requests (e.g., *şeyi veriyim mi?/shall I give this thing?*, extract 10c) and the presentation of their suggestion just as an example or a possibility for the group to consider (e.g., *falán gibi/something like that*, extract 2c) or displaying a stance of certainty towards them through, for instance, their positive claims and assertions about their suggestion in declaratively-formatted turns (extract 4c). By downgrading their suggestions, the participants do not implicate epistemic priority as the first provider of the suggestion (Heritage & Raymond, 2005) but instead invite collaborative thought on the issue and solicit mutual agreement, sharing the responsibility for the decision. In this way, they pave the way for other alternatives and for a careful thought and decision-making on the offered suggestions. On the other hand, the participants’ subsequent display of certainty with their suggestions through a stronger evaluative stance in extracts 2c and 10c (e.g., *changed with it bi garip oluyo/sounds weird*, extract 2c; *o zaman °counsellor mantıklı°/then “counsellor” is plausible*, extract 10c) leads to the acceptance of the suggestion by the group.

In addition, in extract 4c, what seems to play an important role in the acceptance of the suggestion (i.e., omitting the verb “*prefer*” in the new distractor option) is the reviewers’ display of a stance of certainty towards it *in the first place* by presenting it in declaratively-formatted turns and imperative forms (e.g., *dersin/say, de/say, yazma/don’t*

*write, at/omit*). As the suggestion-makers, GE and DU initially tell SA what to include in the option (i.e., *buy*) and then instruct her what not to write (i.e., *prefer*) to clarify which word SA should leave out, bringing the unnecessary word “*prefer*” to the forefront of the group’s attention. This shows that the participants make an attempt to ensure shared understanding about the suggestion by extending their turns and making further clarification, which is another indicator of their display of a stance of certainty towards their suggestion. Finally, another interactional resource that was found to be employed in the suggestion-making sequence in phase RC is the use of a preface (extracts 2c, 10c), which projects the announcement of an upcoming suggestion, drawing the other participants’ attention to it and thus, initiates the suggestion-making sequence (e.g., *şey diyebiliriz/we can say this*, extract 2c; *şeyi veri+yim mi?/shall I give this thing?*, extract 10c).

Before ending phase RC and thus, the IR interaction, the participants might also make a final comment on the revised item (extracts 2c, 3c, 4c, 10c). In doing this, they refer to the main issue addressed in phase PI in order to establish common ground on the underlying reason for the revision of the item at the point when the participants shift their focus to a different issue while reviewing it and to make sure that the problem is solved and is clearly understood by the group (extracts 2c, 10c). They also offer a positive assessment of the change already made in the item as in extract 3c (i.e., *°güzel oldu bu da°/this has also become nice*) and going beyond the main issue, they comment on the difficulty level of the revised item (i.e., *en zor zoru- (.) soru da bu olur bence/I think this would also be the most difficult question*) and the plausibility of the distractor in question (i.e., *çünkü olabilecek tek (.) er şey (.) bakacak incelicek/because it is the only thing that can be possible, s/he will look and examine*) as in extract 4c.

Finally, the findings show that throughout the phase of RC, the participants often go beyond the use of an agreement or a compliance token and display mutual agreement and understanding through certain interactional resources. This is important in establishing shared understanding and making the group’s stance clearer so that the change can be overtly confirmed and a joint decision-making process can be followed. As Table 22 illustrates, the participants display mutual agreement with and understanding of the change

in the revised item by (1) collaboratively providing additional account and evidence for the change in the revised item or for a challenged suggestion as in extract 10c, (2) using a repetitional response that overtly confirms the change (Heritage & Raymond, 2012), (3) using a modified repeat (Kim, 2002; Stivers, 2005) of a claim made about the revised item, which displays and demonstrates their shared understanding of the matter at hand, (4) articulating the marked confirmation “*tabi/of course*” (Stivers, 2011), which not only exhibits their agreement with the change but also displays a stance of certainty towards the revised item, resulting in the acknowledgement of the change in the item, (5) offering a (re)formulation of the change suggested in the prior turn (Heritage & Watson, 1979) while marking it on the exam paper, which not only shows one’s uptake of the offered suggestion but also ensures correct and shared understanding of the change so that it can be noted correctly on the exam paper, (6) providing a candidate explanation for the new change suggested in phase RC as a way of demonstrating their shared understanding of the underlying reason for the suggested change and hence, displaying their acceptance of it, (7) offering a positive assessment of the change, which makes their stance clearer and reconfirms the change, (8) responding to a confirmation request with a full-form recognitional reference as seen in extract 1c where BA and DE respond to EK’s question in disalignment through a non-conforming response (Raymond, 2003) by going beyond the terms and constraints of her question and using a full-form recognitional reference (i.e., *meat*) to refer to the accepted alternative and to ensure shared understanding about it, and (9) employing an anticipatory completion (Lerner, 1996a) as ME does in extract 10c in the form of a candidate evaluation, which enables the participants to discuss the addressed issues jointly and approach them as assumed knowledge. As a result, those interactional resources overtly confirm the revised item or the issue raised about it and eventually lead to the final decision through a collaborative process, ending the IR interaction.

**Table 22. Interactional Resources for Displaying Mutual Agreement and Understanding in Phase RC**

| No | Display of Mutual Agreement                                       | Extract No      | Examples                                                                                                                                                                                                                                                                                         |
|----|-------------------------------------------------------------------|-----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Additional account/evidence                                       | 1c, 3c, 4c, 10c | ara bağlaç olması lazım/ <i>it should be a coordinating conjunction</i> (line 19, extract 3c)<br>da↑ha çok alışveriş yapsın diyo ama expensive olması söz konusu değil/ <i>it says s/he should do more shopping but its expensiveness is not the point in question</i> (lines 24-25, extract 4c) |
| 2  | Repetitional response                                             | 1c, 3c, 4c      | oldu/ <i>it is all right</i> (line 4, extract 3c)                                                                                                                                                                                                                                                |
| 3  | Modified repeat of a claim                                        | 1c, 4c          | bahsetmiyo/ <i>it does not mention it</i> (line 43, extract 1c)                                                                                                                                                                                                                                  |
| 4  | Marked confirmation “tabi/of course”                              | 1c, 3c          | tabi tabi/ <i>of course</i> (line 43, extract 1c)                                                                                                                                                                                                                                                |
| 5  | (Re)formulation of suggestion turn                                | 2c, 4c          | with yerine o zaman using diyelim/ <i>we shall say “using” instead of “with”</i> (line 12, extract 2c)                                                                                                                                                                                           |
| 6  | Candidate explanation for the change                              | 2c, 10c         | o (.) cümleye uya- uyarlamak için/ <i>to adapt it for that sentence</i> (line 14, extract 2c)                                                                                                                                                                                                    |
| 7  | Positive assessment of the change                                 | 2c, 3c          | °güzel oldu bu da°/ <i>this has also become nice</i> (line 24, extract 3c)                                                                                                                                                                                                                       |
| 8  | Responding to a question with a full-form recognitional reference | 1c              | meat'li daha iyi/ <i>it is better with “meat”</i> (line 5)                                                                                                                                                                                                                                       |
| 9  | Anticipatory completion                                           | 10c             | iki kız kardeş oldu/ <i>it turned out to have two sisters</i> (line 7)                                                                                                                                                                                                                           |

The analysis of the talk-in-interaction in phase RC shows that an orientation is displayed to the *institutional role distribution of the participants*. As revealed in this study, the interactional actions emerging in this phase show that the testers and reviewers have certain roles and responsibilities. To begin with the interactional indicators for tester roles, as illustrated in Table 23, it is seen that as the writers of the test items who are accountable for the test, as in phase SC, the testers check and ensure correct understanding of the change that is either already accepted by the group in phase SC or newly produced in phase RC. In doing

this, they, for instance, offer a candidate understanding (Antaki, 2012; Heritage, 1984b) of the change as seen in extract 1c or provide a (re)formulation of the prior suggestion turn (Heritage & Watson, 1979) by explicitly articulating the kind of change that is suggested and accepted while marking it down on the test material as in extracts 2c, 4c and 10c. In this way, they direct the trajectory of the interaction towards the final decision and ensure mutual agreement in the group with respect to the change before completing the revision of the test item. Such interactional resources function as a final check for the item in and through talk and prevent an incorrect noting of the change on the test material, orienting to the testers' responsibility for producing the end product and for being accountable for the test. In a similar vein, the testers' non-verbal behaviour of marking the decided new change on the test material orients to their responsibility and obligation to finalize the test item and make it ready for administration.

**Table 23. Summary of Interactional Indicators for Tester and Reviewer Roles and Responsibilities in Phase RC**

| No                                  | Tester Roles and Responsibilities                     | Extract No  | Examples                                                                                                      |
|-------------------------------------|-------------------------------------------------------|-------------|---------------------------------------------------------------------------------------------------------------|
| 1                                   | Checking correct understanding of the change through; |             |                                                                                                               |
|                                     | candidate understanding,                              | 1c          | <i>meat</i> (.) <i>dedik di mi?</i> =milk <i>değil?</i> /we said "meat" right? not "milk" (line 12)           |
|                                     | (re)formulation                                       | 2c, 4c      | with <i>yerine o zaman</i> using <i>diyelim</i> /we shall say "using" instead of "with" (line 12, extract 2c) |
| 2                                   | Noting down the new change on the test material       | 2c, 4c, 10c | <i>prefer'ü atıyım</i> /let me omit "prefer" (line 8, extract 4c)                                             |
|                                     |                                                       |             | Line 35 (extract 10c)                                                                                         |
| Reviewer Roles and Responsibilities |                                                       |             |                                                                                                               |
| Before the IR Session               |                                                       |             |                                                                                                               |
| 3                                   | Reading the story for the reviewing of the story exam | 10c         | <i>bet mi vardı sonunda betty mi vardı</i> /was there "bet" in its end was there "betty" (lines 20, 22)       |



**Table 23 (continued)**

| <b>During the IR Session</b> |                                                                      |                                                                                                                                                          |
|------------------------------|----------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| 4                            | Evaluating the change from students' and teaching perspectives       | 1c, 2c, 4c, 10c                                                                                                                                          |
|                              |                                                                      | çok (.) öğrettik ama/we have taught it a lot but (line 28, extract 1c)<br>bu kız kardeşleri biliyorlar/they know these two sisters (line 8, extract 10c) |
| 5                            | Providing further clarification of the problem announced in phase PI | 2c, 10c                                                                                                                                                  |
|                              |                                                                      | önemli olan orda (.) şeyi-bilmeleri past tense'ini/the important thing there is its thing for them to know its past tense (line 16, 18, 22, extract 2c)  |

As for the reviewer roles in phase RC, as Table 23 demonstrates, in and through the talk-in-interaction, the reviewers orient to some of the responsibilities that they should carry out both before and during the IR session as in the case of the other phases examined. To begin with the former (i.e., *before* the IR session), in the case of the reviewing of the story exams, the reviewers are required to read the story in question before coming to the IR session so that they can take the exam as a student would and critically review it. This is made interactionally evident in extract 10c as the reviewer YU tries to recall names from the story (See Table 23.3) and displays knowledge of the suggested names by, for instance, providing account for “counsellor” and challenging “judge referson”. This demonstrates that he has fulfilled his responsibility to read the story carefully before coming to the IR session for the story exam, orienting to his K+ status about the issue (Heritage, 2013).

As for the responsibilities carried out *during* the IR session in the phase of RC, the reviewers mainly show an orientation to their roles as teachers by evaluating the revised item or an offered suggestion from the students' and teaching perspectives. This implies the reviewers' epistemic right, responsibility and obligation (Heritage, 2013) to review the change with consideration of the context of learning as well as the target group of students' assumed L2 knowledge and abilities, which is a domain of information that the teachers have a greater epistemic access to. To illustrate, in extracts 2c, 4c and 10c, the reviewers make a positive comment on the revised item by referring to the students with the third-person plural pronoun and indicating the fulfilment of some testing principles from their perspectives. While the reviewer YE clarifies the previously announced problem (i.e., “test what you

teach” principle) in terms of the students’ knowledge of the verb given in the blank to be transformed into its past tense form in extract 2c, the reviewer DU remarks on the difficulty level of the revised item in extract 4c considering the students’ expected ability to carefully examine the reading text. In addition, as seen in extract 1c, the reviewers also challenge a suggestion offered in phase RC from the students’ and teaching perspectives. In extract 1c, for instance, the reviewer DE challenges the suggestion “*dairy products*” in terms of the students’ knowledge of the word “*dairy*” and BA counter-challenges DE by making reference to past instructional events based on a shared curriculum as evident in her use of an inclusive language for their teaching practice (i.e., *öğrettik/we have taught*) (Can Daşkın & Hatipoğlu, 2019; You, 2015). As a result, those interactional resources indicate that the reviewers have the role of reviewing and evaluating an item in terms of the testing principles that mainly concern the students’ L2 knowledge and abilities as well as the syllabus and the instructional events.

Secondly, as Table 23.5 illustrates, the reviewers display an orientation to their roles as problematizers of the item by providing further clarification of the problem they announced in phase PI and that led to the change process in the item. This signals their responsibility to remind the IR group of the main issue and to establish common ground on the underlying reason for the revision of the item at the point when the participants shift their focus to something else (i.e., the grammatical accuracy of the revised sentence in extract 2c and the plausibility of the options/extra character names in extract 10c). Hence, as the problematizers of the item, they make sure that the problem is addressed and is clearly understood by the group. For instance, in extract 2c, the reviewer YE refers to the violated testing principle, reminding the group that what actually matters is the students’ knowledge of the past tense of the verb that is tested. By bringing to the forefront the underlying reason for the change process in the item and clarifying it, she makes sure that the main issue is not overlooked as they shift their focus to the grammaticality of the sentence and ensures shared understanding with regard to their primary aim in revising the item.

Despite an orientation to the institutional role distribution of the participants, it is seen that they share the responsibility for reviewing the change in the revised item and are equally engaged in the phase of RC. Whether it is the reviewer or the tester, they all have a contribution to make to the item change reviewing process. They all, for instance, employ interactional actions such as reviewing the change with reference to different testing

principles by, for example, reading out the target sentence with the insertion of the new distractor “*whereas*” into it as in extract 3c, soliciting the group’s confirmation for the change, challenging the change in the revised item, giving account for the change, suggesting a new change and evaluating it together. Based on this, they are often found to follow a collaborative process throughout the talk-in-interaction in finalizing the item by seeking each others’ confirmation for the change and thus, inviting the group to recheck the item, downgrading their suggestions, and displaying mutual agreement and shared understanding through such interactional resources as anticipatory completions, additional accounts, modified repeats and candidate explanations for the change. In relation to this, the analysis has revealed that though the participants happen to deploy various interactional actions, they sometimes assume certain situated roles depending on the nature of the interaction that unfolds in each extract and make a different contribution to the item reviewing process. For instance, in extract 2c, while the tester SA raises the problem in the grammatical structure of the revised sentence through her confirmation request, the reviewer CI suggests a change to overcome this problem (i.e., replacing the preposition “*with*” in the sentence by “*using*”) and the other reviewer YE offers a candidate explanation for the suggested change and then provides a further clarification of the main issue (i.e., students’ knowledge of the verb given in the blank). Thus, all the participants actively and critically think about the revised item in and through talk and follow a joint decision-making process. This is because the participants assume equal epistemic state and authority since they have equal epistemic access to the available resources (e.g., reading text), are in K+ position in terms of their English language expertise, and share a common ground in terms of the testing principles for language test construction (Heritage, 2013).

The participants’ equal active engagement in the phase of RC is also evident in their gaze behaviour. It is seen that throughout the interaction, they continuously look at their exam papers only sometimes directing their gazes at each other for a short time. However, this does not indicate disengagement; on the contrary, it signals that the participants actively think about the change made in the item in such a competing task-oriented activity (Goodwin, 1981; Rossano, 2013). In addition, the participants’ non-verbal actions such as scanning through the text, doing silent reading, turning over the page to look at the reading text, searching through the story book, and producing thinking face further demonstrate their active engagement in the phase of RC as well as their shared responsibility for the revised item and their ownership of the test as the shared product, the outcome of which is important

for all of them. Finally, the participants' use of an inclusive language throughout the interaction in phase RC also represents their shared responsibility for the change process in the item and shows that the revised test is the product of the whole group (extracts 1c, 2c, 10c).



## **CHAPTER 5**

### **DISCUSSION**

#### **5.0. Introduction**

This chapter discusses the findings in terms of the overall structural organization of language test IR interactions in the case of problematized items including the phases that constitute those interactions and the sequential environment of each phase. In the following sections, the testing principles oriented to and the institutional role distribution of the participants displayed in the phases of language test IR interactions are explained. This is followed by a discussion of findings in terms of their implications for language test construction, EFL teachers' professional competence and for L2 teacher education.

#### **5.1. Overall Structural Organization of Language Test IR Interactions in the Case of Problematized Items**

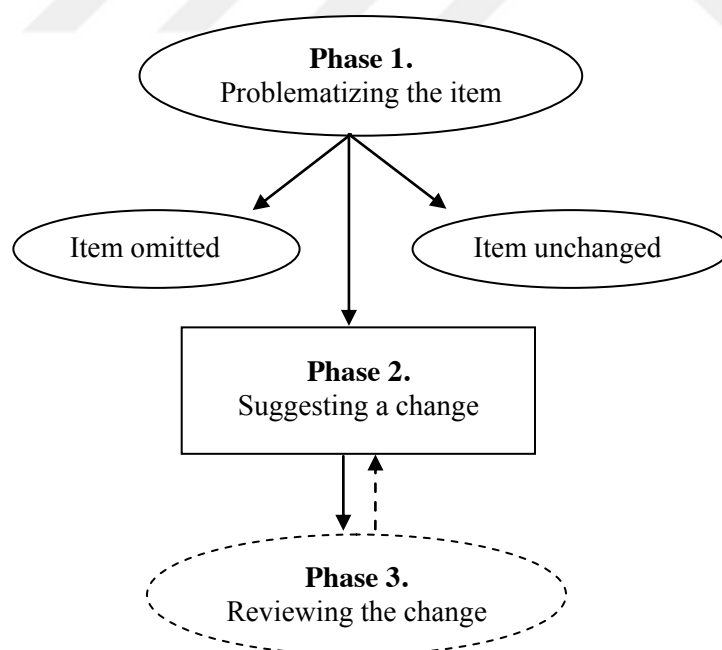
In this section, the phases that have emerged in the language test IR interactions examined and the sequential environment of each phase are explained respectively.

##### **5.1.1. The Phases of Language Test IR Interactions**

The analysis of language test IR interactions in instances of problematized items showed that they consist of three recurring phases, namely (1) problematizing the item (PI), (2) suggesting a change (SC), and (3) reviewing the change (RC). As Figure 6 illustrates, these emerging phases make up the overall structural organization of the IR interactions where an item is evaluated negatively and is disapproved by the reviewers. The first phase is about the participants' announcement of a problem they have identified in the item and their discussion of the problem. This phase ends when the group reaches a consensus about the raised issue and if the group agrees on the problem, the second phase of SC is initiated as some subsequent action is made relevant. This phase includes suggestion-making sequence(s) so as to revise the item that is problematized in phase PI and to jointly decide on

a change. Once the group agrees on a change, the tester notes it on her exam paper. If the participants this time orient to the revised item to check its appropriacy after undergoing a change, the final phase of RC takes place and the IR interaction eventually ends when they fully confirm the change and finalize the item.

Overall structural organization is one of the dimensions of interaction proposed to examine the institutional character of a talk-in-interaction and to show how participants orient to the institutional context of their talk as they organize their conduct (Drew & Heritage, 1992; Heritage, 1997, 2005; Heritage & Clayman, 2010). As Drew and Heritage (1992) state, many kinds of institutional interactions are “characteristically organized into a standard ‘shape’ or order of phases” that are often task-related and that take place recurrently across ranges of instances, persons and so on with a joint orientation towards them (p.43). The phases this study has identified in the IR interactions constitute an achieved, situated and shared orderliness established by the teachers for managing their routine task (i.e., reviewing of language test items) in and through interaction (Drew & Heritage, 1992). They provide an interactional coherence and facilitate the IR process.



\*See Appendix C for the meaning of flowchart symbols

**Figure 6. Overall Structural Organization of Language Test IR interactions in the Case of Problematized Items**

The three phases stem from each other and thus, are dependent on the occurrence of a previous phase. In this respect, the phase of PI constitutes the cornerstone of the talk-in-interaction in the reviewing of problematized items and the phase of SC is made conditionally relevant next only if the participants agree with the problem so that the item can be revised to solve this problem. The analysis shows that in most of the instances of PI, the phase of SC is projected as the relevant next action upon the participants' acceptance of the problem (extracts 1a, 2a, 5a, 7a, 8a, 9a, 10a, 12a). It is commonly seen that when closing the sequence in phase PI, the participants request and seek a suggestion for a change in the item (e.g., *oraya başka bişey ne olabilir/what else can be put there*, extract 7a), which is embodied in extracts 5a and 7a with their orientation to the test material and other relevant sources (e.g., textbook) in search of a suggestion for a change in the item. This not only projects continuity in the IR interaction but also conveys the participants' agreement with the problem as well as their willingness to change the item and move onto the next phase of SC in the IR interaction. As for the phase of RC, it can only take place on condition that a change is made in the item in the prior phase because phase RC is directly related to the reviewing of this new change or the newly revised item as a whole. In this respect, there is a move from the item reviewing format in phase PI to the item revision format in phase SC and then a return to the reviewing format again in phase RC. This means that each stage in the overall structural organization of IR interaction pursues a specific goal or a sub-goal and is jointly constructed by the participants in the process of successfully completing the IR task in and through interaction (Heritage, 2005; Heritage & Clayman, 2010). Thus, overall structural organization enables both progressivity of an activity through the completion of its components and the projectability of the actions, which are organized relative to each other and which have a reflexive relationship (Robinson, 2013).

This study has also shown that the discovered phases in IR interactions are locally managed and thus, are subject to variations with certain stages being optional in some instances depending on a range of contingencies (Drew & Heritage, 1992). This variation is demonstrated in Figure 6 where it is seen that in some instances, the reviewing interaction may end just in the first phase without proceeding with the rest of the phases although it is a rare situation. This occurs when the group decides to omit the item as a result of problematizing it (e.g., extract 12a) or to keep it unchanged when the group challenges and disagrees with the presented problem (e.g., extract 11a). In a similar vein, the last phase (i.e.,

RC) is also found to be optional in the data since the changes made in an item in phase SC are not always subject to further reviewing by the group. Thus, in those instances, the interaction ends once the problematized item is revised and the decision with respect to the change is confirmed in the second phase. This is the reason why phase RC is not projected as the relevant next action when closing the sequence in phase SC in the same way phase SC is made conditionally relevant next in phase PI due to the obligation to solve the announced problem in the items and revise them accordingly (Heritage, 2013). RC is launched only on one's own initiative without being projected, appearing as an extension to the other phases.

With respect to the variations in the phases of IR interactions, the findings further showed that the previously completed phases can be reinstated and the phases can have an interlocking organization. This is seen in phase RC when the newly revised item is challenged and a further modification is required in the item. This makes the second phase of SC conditionally relevant next so that a new suggestion and decision making process can be operationalized. Thus, a transition back to phase SC can occur while phase RC is being employed. Even in the same phase, it is seen that the pattern can be repeated. In the last phase of RC, for instance, the participants may start over the item change reviewing process to do a final check from different perspectives. Likewise, phase SC is repeated when a suggestion is challenged and rejected by the group members.

As Heritage (2005) points out, the discovery of an order of phases does not mean that each of those phases will always take place in every institutional interaction of a particular type and "the parties may well reopen sections and reinstate task orientations that they had previously treated as complete" (p.122). As a result, rather than finding invariance in the order of phases, the purpose of describing overall structural organization is to "identify task orientations which the participants routinely co-construct in routine ways" and hence, it is not a "framework -fixed once and for all- to fit data into" (Heritage, 1997, p.229). The participants' orientations to the goal and task of language test IR practice in and through the moment-to-moment talk may, thus, lead the interaction in different directions depending on the outcomes of each phase. The variation observed in the overall structural organization of language test IR interactions, however, still point to the systematicity involved in the management of those interactions and how the IR activity is "talked into being" through particular phases of talk and routine task orientations (Heritage, 1984a, p.290).



The identified overall structural organization demonstrates the institutionality of language test IR interactions between teachers in educational settings. The interactional architecture is related to the goal of IR practice, which constrains the actions and participants' contributions (Drew & Heritage, 1992; Gardner, 2004; Heritage, 2005; Heritage & Clayman, 2010; Seedhouse, 2011; Walsh, 2011). Since the goal is for the teachers to review the test items in terms of the relevance and usefulness of their content and format for the intended purpose, they present an evaluation of test items and as part of their institutional responsibilities in this task-oriented practice, they go beyond this to revise the problematized items in a way that everyone agrees with the final product. Therefore, the three phases represent the distinctive, specialized, restricted and recurrent character of language test IR interactions taking place in an institutional context (Drew & Heritage, 1992; Heritage & Clayman, 2010; Zimmerman & Boden, 1991). They show how participants shape their conduct in and through talk when a problem emerges in an item and demonstrably make relevant in interaction the institution-specific tasks and actions that they are expected to perform in IR sessions (i.e., announcing a problem, suggesting and deciding on a change, reviewing the change).

Contributing to a corpus of 'workplace studies' which focus on the interplay of work, interaction and the use of tools and technologies in a range of organizational environments (Heath & Luff, 2013), the analysis of language test IR interactions in terms of their overall structural organization unveils how an institution-specific artifact, that is the test material being reviewed structures the ongoing activity and undergoes change within the collaborative production and interactional accomplishment of the IR activity. The items in the test material are made available for evaluation and it is those co-present test items that are problematized, revised and reviewed again after being modified, which are the central actions that constitute the phases in IR interactions. Therefore, evaluative practice is integral to language test IR interaction, functioning as a resource for achieving its institutionality (Lindström & Mondada, 2009). The evaluative practice not only shapes the activity and the institutionality of the educational context but it is also adjusted to the specificities of this context, leading to the emergence of the three phases (Lindström & Mondada, 2009).

The phases as the underlying interactional pattern of IR interactions demonstrate the relevance and procedural consequentiality of these interactions for the task at hand, further determining the institutionality of talk (Schegloff, 1992a). That is, the participants orient to

and establish the relevancies of IR practice by problematizing the test items, suggesting changes to revise them and reviewing the changes made in the item. The participants' orientation to the institutional context and the goal of IR practice is also procedurally consequential to the talk as it shapes the institutional character and trajectory of the talk-in-interaction and affects its outcomes (e.g., deciding whether the item is problematic, is to be omitted or changed). Thus, in this study, it is seen that IR interactions have causes and consequences (Heritage & Clayman, 2010) since the test items are not only evaluated and problematized as part of the requirements in test validation process but they also go through a decision-making and change process and are subject to constant checking, negotiation and compromise as illustrated by the emerging three phases.

### **5.1.2. The Sequential Environment of Each Phase**

In this section, the micro-analytic details in the sequential environment of the three phases identified in language test IR interactions with respect to problematized items are discussed. Each phase is explained in separate sub-sections.

#### **5.1.2.1. Problematizing the Item (PI)**

To begin with the first phase of PI, the analysed extracts in this study demonstrate that problematizing a language test item in an IR interaction is not simply about stating the identified problem but rather it involves some interactional actions and resources that display an orientation to the violated testing principle and lead to the collaborative investigation, recognition, explanation and acceptance of the problem in and through talk. As a result, the raised issue with regard to the language test item is critically evaluated by the co-participants through talk-in-interaction and a better understanding of the problem is eventually reached. This is an important step either for the revision and improvement of the item in the following phases of the IR interaction or for the decision to omit the item from the test as in extract 12a. It is also seen that such a collaborative procedure even leads to the redesign and redefinition of the presented problem as in extract 3a and shapes the act of problematizing the item during the talk-in-interaction through constant checking, negotiation and compromise. Even when the problem is directly accepted as in extract 8a without being subject to further collaborative inquiry and search for evidence by the group, the problem is found to be tailored to ensure mutual agreement in the group (e.g., presenting the problem as

the product of shared understanding among teachers). Furthermore, as illustrated in extract 11a, the findings show that as a result of the co-participants' implementation of certain interactional actions, a language test item that initially seems problematic may be decided to be kept unchanged once the group agrees that the item, in fact, causes no problem and that there is no threat to its validity.

The overall structural organization of phase PI was illustrated in a model in Figure 3 (See Section 4.1.1) where the interactional actions and their systematic sequential organization in the collaborative production and sequential accomplishment of this phase are apparent. Table 24 shows a summary of the recurring interactional actions employed by the participants throughout the phase of PI in the language test IR interactions. In this way, a vivid picture can be drawn of how the participants orient to the violated testing principles in the reviewed items and shape the process of problematizing the item. The interactional actions presented in Table 24 occur subsequently except that the acts of challenging a claim and displaying mutual agreement may also take place alternatively in response to the announcement of the problem in the test item and that the act of performing mitigating actions occurs simultaneously with the acts of giving account for the problem and challenging a claim. In this section, those interactional actions are explained and discussed respectively as given in Table 24.

**Table 24. Recurring Interactional Actions in Phase PI**

| No | Interactional Actions                                                         | Extract No                                    |
|----|-------------------------------------------------------------------------------|-----------------------------------------------|
| 1  | Preliminary inquiry into the problem                                          | 2a, 4a, 5a, 9a, 10a, 12a                      |
| 2  | Giving account and evidence for the problem                                   | All extracts                                  |
| 3  | Challenging a claim                                                           | 3a, 4a, 9a, 10a, 11a, 12a                     |
| 4  | Displaying mutual agreement and understanding through a collaborative process | 1a, 2a, 3a, 4a, 5a, 6a, 7a, 9a, 10a, 11a, 12a |
| 5  | Performing mitigating actions                                                 | 4a, 5a, 6a, 7a, 10a, 12a                      |

As illustrated in Figure 3 in Section 4.1.1, the reviewers tend to do some preliminary work and project the upcoming negative evaluation of the item before announcing the problem they have identified in the item. In doing this, they were found to launch a preliminary inquiry into the problem in a pre-expansion sequence to establish the necessary pre-conditions and lay the groundwork for the statement of the problem (Pomerantz & Heritage,

2013; Schegloff, 2007; Stivers, 2013). The reviewer who problematizes the item initiates a preliminary inquiry (Pomerantz & Heritage, 2013) into the problem by (1) seeking concrete evidence for it through a question directed at the other participants with reference to past instructional events based on a shared curriculum (Can Daşkın, 2017; Can Daşkın & Hatipoğlu, 2019; You, 2015) (extracts 2a, 12a) or to shared sources as external sources of authority (Jacknick, 2013) (e.g., the story book, extract 10a), (2) soliciting confirmation for the answer of the item (extracts 4a, 5a), or by (3) expressing uncertainty about the correct answer (extract 9a) (See Table 6 in Section 4.1.1). These interactional resources in the preliminary inquiry not only project a problem in the item but also lead to a process of collaborative inquiry about the issue raised with respect to the item as the participants orient to the target point in the test material (e.g., reading the other option that could possibly be correct, extracts 9a, 11a) or the textbook (e.g., checking where the tested word “*ambitious*” appeared, extract 12a).

The questions addressed to the participants with reference to their past instructional events based on a shared curriculum (Can Daşkın, 2017; Can Daşkın & Hatipoğlu, 2019; You, 2015) (e.g., *hatırlıyor musun derslerde böyle hani cover'ın (.) past tense'i geçti mi?/do you remember if ever in classes like you know the past tense of "cover" was encountered?*, extract 2a) or to shared sources results in the identification of the sought-after evidence for the projected problem in and through talk (e.g., finding out that the verb “*cover*” in the item was not studied in class, extract 2a). Through their questions which appear as either “request for information” (extracts 2a, 12a) or “request for verification” (extract 10a) (Goodwin, 1987), the reviewers display uncertainty about their recall of past instructional events and the content of already studied sources as further evident in, for instance, YE’s thinking face and claim of inability to recall (*hatırlayamadım/I could not remember*) in extract 12a, and seek the other participants’ help in their search for evidence (Can Daşkın, 2017; Goodwin, 1987). As Wasson (2016) found in her study on negotiation in business meetings, “speakers may set up a negative assessment by preceding it with a request for information that is expected to provide support for the subsequent negative assessment” (p.389). Therefore, they do not position themselves as occupying a K- epistemic status but as lacking access to their K+ epistemic status (Heritage, 2012, 2013). This lack of epistemic access is then resolved with the other participants’ further references to past instructional events and to the specified content of the shared sources that they can recall to locate the evidence

sought for the projected problem. They give negative responses to the reviewers' questions, which mark the item as problematic and act as a go-ahead response for the announcement of the problem (Schegloff, 2007). As a result, the evidence does not happen to be pre-determined and the problem is not imposed on the participants but rather they all get engaged in the evaluation of the item, identifying and justifying the evidence together. This makes the upcoming announcement of the problem in the base FPP more likely to be accepted by the group and builds the main action on a strong ground (Pomerantz & Heritage, 2013).

Reference to past events was also found to be employed in pre-sequences in Balaman (2019), Can Daşkın and Hatipoğlu (2019) and You's (2015) studies as an interactional resource to signal and set the ground for a forthcoming base action. According to Can Daşkın and Hatipoğlu (2019), who have examined reference to past learning events in L2 classroom interactions, such a phenomenon "confirms what is shared and establishes common ground, continuity and temporality" (p.16). Therefore, it can be said that in the present study, by referring to past instructional events based on a shared curriculum and/or to shared epistemic sources such as the story book or the textbook that are assumed to have been previously accessed and that are still available in the participation framework, teachers in IR interactions orient to the temporal relationship between teaching and testing as well as the interconnectedness and continuity between their institutional sources (e.g., between the story book/textbook and the test material). This allows the group to discuss the matter at hand based on a common ground as members of an epistemic community in the institution (Heritage, 2013). This, in turn, ensures progressivity of the interaction in the base sequence when the problem is announced since the evidence required to explain the problem is already located in past instructional events or shared sources by the participants' responses to the preliminary inquiry in the pre-expansion sequence (Can Daşkın, 2017; Can Daşkın & Hatipoğlu, 2019). Reference to past events and available sources, hence, establishes the necessary pre-conditions and helps avoid dispreferred responses to the base sequence (Schegloff, 2007).

Additionally, as in You's (2015) study, reference to past instructional events or sources that are based on a shared curriculum can be said to do claim-backing because the evidence sought for the projected problem in the item is confirmed based on shared experience between the participants and is then used to explain the problem in the item. However,

different from You's study where reference to past events was employed to account for a prior claim within one's own turn, in this study, it has a prospective function rather than a retrospective one as it takes place in a pre-sequence and presents evidence to back up an upcoming claim in the subsequent base sequence.

As for the other interactional resources used to initiate preliminary inquiry in the pre-sequence, the reviewers' confirmation request for the correct answer of the item as well as their expression of uncertainty about the correct answer allow the participants to present account and evidence for the expected answer (extracts 4a, 5a, 9a). To illustrate, in extract 4a, upon one of the reviewers DU's request for confirmation for the correct answer (i.e., *dört "d" di mi/four is "d" isn't it?*), the other reviewer GE shows the part of the reading text where the students are expected to derive the answer from. She also makes reference to past instructional events and shared textbooks where the word key to the correct answer is assumed to have been studied, making sure that it does not pose a threat to the validity of the item. In a similar vein, the testers not only announce the expected answer or (dis)confirm the reviewers' answers but also provide a pre-emptive explanation of the item writing process and their aim in the design of the item (extracts 5a, 9a), which further accounts for the expected answer. Consequently, before the problem is actually announced, the reviewers' confirmation requests for the correct answer and their expression of uncertainty about it lead the trajectory of the IR interaction in such a way that all the participants end up re-evaluating the item, considering what could go wrong with it and sharing their opinions about it through the talk-in-interaction. Through a preliminary inquiry, the reviewer who intends to problematize the item learns what the expected answer is, what was intended in the design of the item and what opinion the other participants hold with regard to the item in a pre-expansion sequence. In this way, shared understanding and common ground is established before proceeding to the base sequence of announcing the problem in the item. This sets up an environment for problematizing the test item and promotes the smooth flow of the IR interaction, gradually shifting the topic to the identified problem.

As a result of the reviewers' preliminary work through questions with reference to past instructional events or shared sources, their confirmation requests for the correct answer and their expression of uncertainty about it, the participants give careful thought to the item in the pre-expansion sequence, which allows the problematizer of the item to confirm the

problem and ensure shared understanding and common ground on which to base the problem safely. In other words, the preliminary inquiry encourages the production of the projected problem as the participants' responses act as a go-ahead response (Schegloff, 2007) and suggest an upcoming acceptance of the problem likely (Pomerantz & Heritage, 2013). In this respect, the preliminary inquiry serves the base sequence of problematizing the item by checking the pre-conditions and laying the groundwork (Schegloff, 2007; Stivers, 2013). In addition, it invites a collaborative investigation and identification of the problem and manages to involve all the co-participants in the process of problematizing the item in and through the talk-in-interaction. Consequently, they collaboratively show an orientation to the testing principle that was violated.

Following the preliminary work, the base FPP of PI is then launched where the problem is overtly announced and explained with evidence (See Table 24.2). It is found that the interactional act of giving account and evidence for the problem is an inevitable part of PI, appearing in all the extracts analysed for this study. However, in such a task-oriented interaction as the IR interaction that includes not just approving but also negatively evaluating test items, there are boundaries in terms of how the accounts should be framed so that the problematizing action can be made relevant to the institutional task and goal (Drew & Heritage, 1992). As summarized in Table 8 in Section 4.1.1, in the base FPP, the reviewers make an attempt to prove the problem by presenting evidence from the available shared sources such as the reading text, textbook, listening notes/script, and the story book as external sources of authority (Jacknick, 2013) as well as the past instructional events (Can Daşkın, 2017; Can Daşkın & Hatipoğlu, 2019; You, 2015) and the students' present state of L2 knowledge and skills. In some cases as in extracts 2a, 10a and 12a, the reviewers announce the problem based on the evidence they have obtained through their preliminary inquiry in the pre-expansion sequence. Furthermore, while providing a justification for the problem in the item, the reviewers at the same time show an orientation to the testing principles for item writing (See Section 5.2). By invoking larger disciplinary principles, they "pronounce the problem as not isolated and idiosyncratic, but violating some fundamental understanding of the profession" (Waring, 2017, p.23). To illustrate, in extract 1a, BA refers to the implausibility of the correct option with her account "*çok (.) böyle (0.4) açık olmuş/it has become too like obvious*" and in extract 3a, the reviewers DE and BA refer to the violation of "only one correct answer" principle as they present evidence from the text.

The reviewers ground their accounts to epistemic authority (Heritage, 2013) by making reference to “external sources of authority” such as the textbook and the reading text in the test material (Jacknick, 2013, p.199) and to their past instructional events, their students’ L2 knowledge and skills as well as the testing principles adopted in the institution. As the central action that characterizes phase PI, such account-giving based on solid evidence and epistemic authority constitutes the actual reviewing part of the interactions and the rest of the interaction centres around the issues raised by those accounts. Therefore, the reviewers display an orientation to the core goal and task of IR and to the particular constraints set by this task (Drew & Heritage, 1992). As Firth (1995) points out, accounts have generally been considered in a retrospective sense to explain unanticipated or untoward behaviour and “to repair or prevent norm violations” (p.200); however, they do more than this and can also be employed in a proactive, creative and problem-solving sense. As in this study, “accounts establish a basis from which organizationally relevant action may be identified, challenged, and discussed” (Firth, 1995, p.199). That is, they “contribute to the overall structure and trajectory of interactional events” by changing “the current circumstances in order to resolve a work-related problem in a creative and mutually acceptable way” (Firth, 1995, p.201). In Baker’s (1997) study, too, the school staff found social-moral order in a system of ‘ticketing rules’ applied in their school by engaging in considerable accounting work and referring to principles of justice and fairness in the sequential and consequential production of their accounts.

The proactive function of accounts means that they can also be challenged in the process of fulfilling work-related tasks (See Table 24.3). Although accounts for the problems in the test items are framed with reference to epistemic authority (Heritage, 2013), concrete evidence and larger disciplinary principles (Waring, 2017), the analysis has shown that they can still be subject to debate and disagreement due to the participants’ equal epistemic status and the presentation of mutually accessible resources (Firth, 1995). In this study, accounts are, thus, treated as contestable and negotiable objects similar to those in a commodity negotiation in Firth’s (1995) study. The negotiability of accounts and thus, the acceptance of the problem is also greatly determined by the kind of account and evidence provided for the problem. As Verkuyten (2000) makes it clear, “not every explanation is equally acceptable and convincing” (p.454).



It is seen that accounts in relation to the texts in the test material as well as the violation of some testing principles are, in some instances, found to be challenged, entailing further discussion and negotiation in the group. This is because the participants (1) come up with different interpretations of the target sentence in the reading text as in extracts 3a and 9a or of the options in the item as in extract 11a (Sharma, 2012), (2) locate the place of the item in different parts of the reading text as in extracts 4a and 11a, or (3) hold different opinions with regard to the violation of the testing principle that is addressed by the reviewers as in extracts 10a and 12a. As a result, the testers challenge the reviewers' claim by (1) presenting counter-evidence from the available shared sources such as the reading text and the story book (e.g., extract 4a), (2) making counter-claims with respect to the violation of the testing principles addressed by the problematizers of the item (e.g., extract 12a), and (3) referring to the item-writing process (e.g., extract 9a) (See Table 10 in Section 4.1.1). In those instances, the testers are made accountable for the announced problem that directly concerns the content of the item they have prepared and thus, they get engaged in the negotiation of the presented accounts.

On the other hand, accounts with reference to past instructional events, the students' L2 knowledge and skills as well as the textbook are directly accepted without being questioned (extracts 2a, 8a, 12a). In those cases, the evidence is directly accepted due to its presentation as a fact and reality-based phenomenon (Verkuyten, 2000) and its situatedness in the reviewers' professional knowledge domain. The reviewers have direct epistemic access to this domain of information as those who actually teach the content of the test and who have greater command of the teaching process. The preliminary inquiry into the past instructional events in extracts 2a and 12a also play a role in the acceptance of the problem in those instances since the evidence is collaboratively identified and justified especially with the involvement of two teachers as reviewers. Consequently, similar to what Verkuyten (2000) found in his study, accounts have different sequential consequences.

Considering the management of disagreement that arises from the participants' act of challenging the announced problem, the findings demonstrate that the process of account-giving (i.e., production of accounts and responses to them with counter-accounts) results in a consensus in the IR group and the resolution of the emerging dispute in and through the talk-in-interaction, which further illustrates the resolution-oriented aspect of accounts (Firth, 1995). Accounts not only make visible the participants' stance of disagreement but also

“provide mutually accessible discourse-based resources that make the progression from ‘conflict’ to substantive agreement both possible and organizationally defensible” (Firth, 1995, p.205). In the data examined for this study, the management of disagreement is achieved locally through (1) further account-giving and elaboration of the problem as a counter-challenge in response to the disagreeing participant in a non-minimal post-expansion sequence, which serves to create better understanding of the problem and makes it acceptable (extracts 3a, 9a), (2) the problematizer’s redesign of a prior account from a different perspective (e.g., the students’ L2 ability) to reverse the displayed preference and elicit a preferred response (i.e., agreement with the problem) (Pomerantz, 1984a; Pomerantz & Heritage, 2013; Sacks, 1987; Schegloff, 2007) by drawing upon disciplinary knowledge that establishes the severity of the problem in response to the displayed resistance (Waring, 2017) (extract 9a), (3) the problematizer’s alignment with the disagreeing participants by downgrading her competence in answering the item (i.e., claiming carelessness and lack of attention) upon the group’s accounts and (counter)-challenges (extract 11a) (Kotthoff, 1993; Sevgi-Sole & Ünalı, 2020; Sharma, 2012), (4) the disagreeing participant’s quick recognition and understanding of the problem before the other participants’ counter-challenge and her use of a ‘peer referencing’ (i.e., *as you said*) to turn her initial disagreement with the announced problem into agreement (extract 4a) different from the use of ‘peer referencing’ to perform a disagreeing action in Waring’s (2001) study, (5) the disagreeing participant’s backdown from her position as a result of no uptake from the other participant by projecting the next phase of SC in an attempt to receive a preferred response (i.e., mutual agreement in the group) and to ensure contiguity in the IR interaction (extract 10a) (Pomerantz, 1984a, 1984b; Pomerantz & Heritage, 2013; Sacks, 1987), and (6) the announcement of other kinds of problems in the item, which the participants agree with (extract 12a).

As a result, in managing and resolving disagreement, some interactional work is involved on the part of both the problematizers of the test items and the disagreeing participants who all exercise cautiousness to achieve consensus in the IR interactions (Wasson, 2000). Through a reversal of opinion or conceding activity (i.e., giving up one’s position) that is employed on the disagreeing participants’ own initiative or that is encouraged by the accounts of the participants, an orientation is made to the consensus-oriented nature of the delicate work of reviewing and problematizing a test item (Sharma, 2012; Wasson, 2000). This shows that while challenging the account for the problem in the item, the participants at the same time

seek concession and display a preference for agreement over a series of turns that address a better understanding of the issue (Pomerantz, 1984a; Sacks, 1987). As Carr (2011) indicates, the process of reviewing test items should be approached collegially whereby decisions are made by consensus without being enforced upon others. Alderson et al. (1995) also point out that item writers should not be too defensive about their items and should be ready to accept the change or abandon the item if it is found problematic, working towards resolving the disagreements.

The accounts or counter-accounts given for the identified problems in the language test items reviewed in this study are found to have implications for decisions about those test items (See “Outcome” in Table 10, Section 4.1.1). This is also acknowledged by Verkuyten (2000) in his study on teachers’ accounts for their students’ school marks, where it was found that decisions for remedial intervention depended on the kinds of accounts that the teachers gave in interpreting their students’ marks. In the majority of instances in the present study, even if the presented account is challenged by the group members, the problem is eventually approved through a process of open dialogue and negotiation and the item is decided to be revised in a new phase. However, it is also seen that as in extract 12a, the reviewers’ construction of their accounts with reference to a variety of issues such as the absence of the tested word in the past instructional events or the textbooks studied so far, the difficulty in understanding the target sentence, and the violation of the testing principle with respect to the independency of the item (i.e., two points are tested subsequently in one single sentence) results in the group’s decision to omit the item as the identified problems leave no room for change in the item.

Furthermore, although preference is displayed for agreement with the announced problem in the test item (Pomerantz, 1984a; Sacks, 1987) and some interactional work is involved in conceding the problem, it is also found that the testers’ act of challenging it can change the course of action in different ways as in extracts 3a and 11a. In extract 3a, further account-giving and rechecking of the problem entailed by the tester’s act of challenging the problem leads to the redesign and redefinition of the already announced problem during the talk-in-interaction. Therefore, the group decides that not two but only one of the other distractor options in the item (i.e., *therefore*) can be possible as a correct answer within the given context. In another case as in extract 11a, the act of challenging this time disproves the problem claimed for the item and thus, the group decides to keep the item unchanged after

agreeing that it, in fact, causes no problem through a process of considerable discussion on the item. In this instance, as a result of an emerging dispute with regard to the correct answer of the item, different from the other extracts, not the problematizer of the item but the participants who challenge the problematizer through account and evidence turn out to be right in their argument for the validity of the item and hence, the group eventually approves the item.

Although consensus is somehow reached through a reversal of opinion (Wasson, 2000) or a conceding activity (Sharma, 2012) in the IR group, such a procedure involving negotiation and discussion of ideas about the test items at the same time illustrates that in such task-oriented workplace interactions where the product is shared, the participants are responsible to exchange any necessary information they possess about the item and share their views in an open dialogue. In this way, they can reach a mutual understanding about the item and determine the course of action accordingly. As Sharma (2012) found in his study, in small group interactions, “despite their oppositional stances that become apparent through their disagreements in talk, the interactants have opportunities to display multiple perspectives while completing the task” (p.4). They then deploy certain pre-conceding activities to exit from disagreements so that some subsequent action can be performed in relation to the task at hand depending on the outcome of the discussion (Sharma, 2012). In this respect, in the context of language test construction, it is seen that contradictory view is not regarded as a threat to but as a source of validity and a way to resolve conflict (Davison, 2004). As Fulcher and Davidson (2007) indicate, during the test validation process, there will be disagreements and alternative explanations that challenge the validity argument and test validation is about investigating those alternatives and successfully eliminating the counter-arguments to explain the evidence and support the use of a test for a given purpose (Fulcher & Davidson, 2007; So, 2014). Thus, such a collaborative way of understanding the problem and considering all the parameters of the problem in and through talk leads to its acceptance “in open discussion, through a process of dialogue and disagreement” (Fulcher & Davidson, 2007, p.21).

With respect to the participants’ display of agreement with the accounts provided for the problems in the test items (See Table 24.4), it is seen that throughout the phase of PI, they often go beyond the use of an agreement token and display mutual agreement and understanding through a collaborative process, ‘sharing’ the task of accounting for the

identified problem (Firth, 1995) and constructing the arguments about the item jointly (Smithson & Diaz, 1996; Van Kruiningen, 2013). In doing this, they use a variety of interactional resources such as full or a partial modified repeat (Kim, 2002; Stivers, 2005), (re)formulation (Baraldi, 2013; Barnes, 2007; Drew, 2003; Franco & Nielsen, 2018; Heritage & Watson, 1979; Van Kruiningen, 2013), additional account or evidence, candidate understanding (Antaki, 2012; Heritage, 1984b), candidate evaluation and explanation, collaborative turn completions and sequences including anticipatory completion (Lerner, 1996a, 1996b, 2004; Sert, 2019; Szczepek, 2000; Van Kruiningen, 2013) and choral co-production (Lerner, 2002), as well as reflection on the item writing process (See Table 11 in Section 4.1.1).

These interactional resources enable the participants to demonstrate their understanding of the matter at hand (Sacks, 1992) and affiliate with the stance taken by the problematizer of the item in a goal-oriented and collaborative fashion (Hayashi, 2013). They also overtly confirm the announced problem and coordinate shared agreement and understanding (Barnes, 2007). Based on this, by means of such interactional resources, as in Van Kruiningen's (2013) study, the participants further elaborate on the problem especially in a non-minimal post-expansion sequence (Schegloff, 2007; Stivers, 2013), which is predominantly characterized by collaborative account and evidence giving from various perspectives (extracts 3a, 4a, 6a, 9a). In this way, the participants reassert and clarify the problem further in a non-minimal post-expansion sequence even when an agreement is displayed with the problem in the base SPP, which normatively does not project an expansion as in extracts 4a and 6a (Schegloff, 2007). As a result, they establish mutual understanding about the item in the IR group and make sure that they have given the item a careful and a critical thought before proceeding to the next phase of the interaction. This indicates that as Firth (1995) found in his study on commodity negotiation, account-giving is not a unilateral but a multilateral act, involving all the participants in the activity of accounting as a way of displaying mutual agreement and understanding. The account is continuously negotiated and systematically expanded through collaborative joint action (Firth, 1995). As in Verkuyten's (2000) study, a combination of accounts grounds the problem on a solid and factual basis and leads to a better understanding of the problem. Furthermore, by means of such interactional resources as modified repeats, the participants do not only show agreement but also claim their epistemic right to confirm an assertion, indexing epistemic authority and control over the matter at hand since they all have equal

epistemic access to the referent situation (e.g., reading text) and are in K+ position in terms of access to their language and testing knowledge as English teachers (Heritage & Raymond, 2005; Stivers, 2005). That is, they assume their share of the accounting task for the legitimacy of the claims made about the item (Firth, 1995).

Overall, it can be concluded that phase PI involves substantial interactional work and is mainly characterized by the production of accounts to explain the identified problem in the language test items and responses to them in the form of further accounts that either challenge (i.e., counter-arguments), counter-challenge, or support them, displaying agreement and understanding. In this respect, the findings show that accounts are the ubiquitous features of PI phase in IR interactions and “are offered, discussed, challenged, elaborated and thus topicalised” by the IR group members (Firth, 1995, p.205). Thus, it becomes apparent that the accounts are managed interactively and contingently in an attempt to secure agreement (Firth, 1995; Verkuyten, 2000). Different from the interpersonal and pragmatic functions of accounts that have been widely examined particularly with respect to the design of dispreferred responses (e.g., Pomerantz, 1984a; Schegloff, 2007), this study has revealed that in institution-specific task-oriented interactions, accounts perform work-related functions and function as problem-solving resources. In doing this, they invoke organizationally relevant information to problematize test items, challenge the announced problem, deal with counter-arguments (i.e., to counter-challenge) or to support and display agreement with the problem so that the IR group can reach a decision about the item and determine the interactional trajectory of the IR interaction (Firth, 1995; Sevgi-Sole & Ünalı, 2020).

Finally, the findings illustrate that the reviewers generally treat the act of problematizing the item as a dispreferred FPP by using some interactional resources such as a preface, direct or an embedded question, hedges, mitigating account, and positive assessment (Pomerantz, 1984a; Pomerantz & Heritage, 2013) (See Table 24.5). In a study conducted by Wasson (2000), negative assessments in American business meetings were also constructed with various forms of mitigation that conveyed delicacy and indirectness. It has been acknowledged in the literature that preference is not only relevant to SPPs but also to FPPs as some FPP types such as criticism of co-participants and requests are dispreferred actions (Liddicoat, 2007; Schegloff, 2007). In a similar vein, in this study, the act of problematizing in the context of IR interactions is constructed as a dispreferred FPP through the deployment

of delaying devices and mitigations (Liddicoat, 2007). Those interactional resources (1) project the upcoming negative evaluation of the item, (2) mitigate and downgrade the disaffiliative impact of the identified problem, (3) encourage collaborative inquiry into the raised issue, and (4) by delaying the announcement of the projected problem, they lead to the testers' pre-emptive account-giving that elaborates on the projected issue and to their candidate evaluation. In this way, the participants who problematize the test item employ both politeness and solidarity strategies, orienting to not only the transactional but also the relational goals in a workplace interaction and thus, taking into account their relationship with colleagues even when they are fully focused on an institutional task (Koester, 2006). The functions performed by the delaying devices and mitigations used in this study in problematizing a test item are discussed below separately.

Firstly, the use of a preface which can also appear in the form of a mitigating account (extracts 5a, 10a) projects an upcoming negative evaluation of the item and draws the other participants' attention to it rather than announcing the problem abruptly, which would be more likely to engender a disaffiliative response (See Table 7 in Section 4.1.1). Secondly, due to their temporal placement, the preface as a delaying device creates opportunities for the testers to do problematizing and accounting pre-emptively based on what is projected (Heritage, 1988). To illustrate, as in extracts 5a and 10a, the reviewers' mitigating accounts in the preface where they give background information about why the item has caused a problem for them does more than projecting the upcoming negative evaluation of the item as they lead to the tester's candidate evaluation of the item as in extract 5a and her pre-emptive account for the expected correct answer as in extract 10a. In relation to this, in some cases in the analysed extracts, the preface is found to be embodied with a non-verbal behaviour such as pointing (Goodwin, 2003) at the problematic word in the item or in the reading text (extracts 5a, 8a), which not only projects the cause of the problem by establishing a shared focus but also prompts the other participants to even anticipate the projected problem and provide a pre-emptive account for it as the tester does in extract 5a. As a result of the testers' pre-emptive attempts to problematize the item, the other participants' projected dispreferred act of announcing the problem turns into a preferred act of noticing by the writer of the item (Schegloff, 2007). Therefore, the preface enables the reviewers to avoid threatening the testers' face by not abruptly offering a negative evaluation of their product and to save the testers' face by making them notice the matter at hand and implicating their epistemic control over the test material which they have prepared (Heritage, 1988). In this respect, the

incorporation of delay features into the design of the problematizing act makes it other-attentive (Heritage, 1988).

Thirdly, the reviewers' accounts (extracts 5a, 10a) and positive assessment (extract 10a) preceding the announcement of the problem and their use of syntactic and lexical mitigating devices such as questions (extracts 4a, 6a, 7a) or hedges (e.g., *sanki/as if*, *belki/maybe*, *galiba/I think*) (extracts 5a, 12a) in the design of the statement of the problem mitigate and downgrade the disaffiliative impact of the identified problem (See Table 9 in Section 4.1.1). The accounts, for instance, downgrade the reviewers' upcoming claim for the item by showing that it may be their own way of approaching the item that might have caused the problem and leaving it to the tester to decide if it could be considered as a serious problem. According to Heritage (1988), accounts are sensitive to issues of face (Brown & Levinson, 1978; Goffman, 1955) and take into consideration self-other relations. Additionally, as in extract 10a, using a positive assessment and incorporating an appreciation component in the context of offering a negative evaluation of a test item enables the reviewers to "attend to the others' viewpoint before attending to what is relevant from their own perspective" (Heritage, 1988, p.136). Therefore, the use of such resources as accounts and appreciation components before launching the FPP avoid any threat to the other participants and to their social relationship and solidarity (Heritage, 1988).

Furthermore, the use of syntactic and lexical mitigating devices such as questions and hedges in the design of the statement of the problem invites everyone in the IR group to think about the raised issue. This paves the way for a collaborative search of evidence for the problem and collaborative inquiry into the raised issue as the participants orient to the test material. In this way, the problematizers mitigate the disaffiliative impact of the announced problem in the item by treating the identified problem as one possible issue for the group to consider and leaving it to them to decide without imposing it. According to Heritage and Raymond (2005), in assessment sequences, being the first to evaluate a matter at hand that is accessible to all the participants present indexes an implicit claim of epistemic priority and thus, first assessors use a variety of resources to downgrade their claims in order to achieve epistemic congruence with the other participants. Hence, it can be stated that by mitigating their claims about the test item, the reviewers problematizing it might not want to implicate epistemic priority about a matter that directly concerns the tester's practice in the construction of the item or that all the participants have equal epistemic access to (e.g., the



reading text, textbook, the distractors in a multiple-choice item, the story book) (Heritage & Raymond, 2005). As Koester (2006) states, minimizing and downtoning are useful strategies on the transactional level by facilitating the task for the group and on the relational level by conveying the message that what the recipient is being asked to understand “is not too much of an infringement on his or her freedom of action” (p.96). Overall, mitigating devices such as accounts, appreciation components, questions and hedges are used in order not to disrupt the relationship that exists in the IR group and so as to minimize the degree of disaffiliation and conflict that might possibly emerge as a result of problematizing the item (Liddicoat, 2007; Seedhouse, 2004).

In this study, disagreements in response to the item problematizing act are also often oriented to as dispreferred responses by the participants (Pomerantz, 1984a; Pomerantz & Heritage, 2013). Considering the disagreement formulations in the analysed extracts, it is seen that in extracts 4a, 9a, and 10a, the testers launch a “pro forma” agreement (Schegloff, 2007) (i.e., an agreement-prefaced disagreement) whereby a minimal agreement that is initially displayed then turns into disagreement as the participants exhibit partial resistance by providing counter-evidence for the problem (McQuade et al., 2018; Pomerantz, 1984a; Pomerantz & Heritage, 2013; Sacks, 1987). Their “pro forma” agreement mitigates the disaffiliative impact of their upcoming dispreferred response (i.e., challenging of the problem) by delaying their projected display of disagreement as a contiguity-breaking practice and framing the disagreeing response “as an exception or modification of an initial, apparently preferred response” (Schegloff, 2007, p.70). This helps the participants save face prior to the dispreferred talk (McQuade et al., 2018).

#### **5.1.2.2. Suggesting a Change (SC)**

SC as the second phase of language test IR interactions is about the participants’ suggestion and decision making process for a change to overcome the identified problem in the test item through a particular interactional pattern which was illustrated in a model in Figure 4 (See Section 4.2.1). As in many joint decision-making activities in workplace meetings, in phase SC, sequences begin with a team member articulating a proposal followed by the other group members’ evaluations of this proposal and end when they reach a consensus decision (Wasson, 2000). The analysed extracts for this study demonstrate that in the phase of SC, a collaborative process of suggestion and decision making is followed via some interactional

actions while seeking a change for the problematized item and the suggestions are not simply listed one by one as distinct entities (Baraldi, 2013). As Table 25 illustrates, there are certain recurring interactional actions that help the group to interactionally deal with the problem identified in the first phase of the IR interaction and that, in fact, facilitate the change process in the problematized language test items that have become subject to revision. As a result of such a sequential process, the IR group eventually decides on a suggested change to be made in the item so that the identified problem is eliminated and the language test item is in line with the testing principles, causing no problem for the target group of students and the syllabus.

**Table 25. Recurring Interactional Actions in Phase SC**

| No | Interactional Actions                                                         | Extract No                          |
|----|-------------------------------------------------------------------------------|-------------------------------------|
| 1  | Suggesting a change                                                           | All extracts                        |
| 2  | Reformulating a suggestion                                                    | 1b, 4b, 5b                          |
| 3  | Giving account for a suggestion                                               | 2b, 4b, 5b, 6b, 7b, 8b, 9b, 10b     |
| 4  | Challenging a suggestion                                                      | 1b, 4b, 7b                          |
| 5  | Displaying mutual agreement and understanding through a collaborative process | 1b, 2b, 4b, 5b, 6b, 7b, 8b, 9b, 10b |

The interactional actions listed in Table 25 occur subsequently except that the acts of challenging a suggestion and displaying mutual agreement may also take place as an alternative to each other in response to an offered suggestion and that the order of the acts of reformulating a suggestion and giving account, which are optional steps, can change. In this section, those interactional actions are explained and discussed respectively as given in Table 25.

To begin with, at the core of phase SC is suggesting a change to solve the announced problem in the language test item. It is this process of suggestion-making that especially shapes the talk-in-interaction in phase SC and demonstrates how the participants interactionally deal with the problem identified in the first phase of the IR interaction with respect to the language test items that are reviewed. The findings have revealed that the interactions can focus on one particular suggestion (extracts 2b, 3b, 5b, 10b), or two (extracts 6b, 8b, 9b, 10b) or multiple suggestions (extracts 1b, 4b, 7b) before deciding on a change in the problematized item depending on the participants' orientations and evaluations

of those suggestions and thus, the trajectory of the interaction in phase SC (See Table 14 in Section 4.2.1). The suggestions do not come as ready-made solutions but are often formulated spontaneously in and through the talk-in-interaction as evident in the participants' elongated hesitation markers, gaps of silences, intra-turn cut-offs and pauses, repetitions, modifications, withdrawal of a suggestion, silent reading, self-repair, word search, and use of a preface that projects their upcoming suggestion accompanied by their non-verbal actions such as producing a thinking face (Goodwin & Goodwin, 1986) and shifting their gaze between the reading text and the item. Even though those suggestions are produced spontaneously in and through talk without too much consideration beforehand, this seems to have interactional implications as they enable the participants to approach the item cautiously with a critical eye and give immediate feedback, leading to the emergence of new ideas and discussions in the course of solving the problem in the item.

The spontaneous formulation of the suggestion in and through talk also makes it possible for the IR group members to reformulate and shape it through a collaborative process involving such interactional resources as collaborative turn completions and sequences as in extracts 1b, 4b, and 5b (See Table 25.2). They make some slight changes and modifications in the new alternative by drawing upon their language expertise before evaluating it in terms of its appropriacy in eliminating the problem in the item and before displaying a stance of (dis)agreement with respect to it. The reformulation or redesign of a suggested alternative is most apparent in extract 5b where the talk-in-interaction centres around the reformulation of one particular time expression suggested by MI as a rephrase of the trouble source "*before*" in the target sentence. Every time the suggested expression is articulated, it becomes subject to small changes, eventually being finalized and decided on. As a result, "*in their early school*" initially articulated just in the process of being formulated is eventually completed and changed into "*in the early years of their education*" through the participants' joint action. Such a process represents how an entry proposal for a textual change is transformed into a joint statement and how shared text production is locally achieved in an institutional meeting "with the text under construction being revised and born step-by-step as the conversation around it unfolds" (Nissi, 2015, p.1).

It is also found that the participants produce their suggestions by downgrading them through hedges, confirmation requests and the presentation of the suggestion as an example (extracts 1b, 2b, 4b, 5b, 6b, 7b, 9b) (See Table 15 in Section 4.2.1). It is revealed that the epistemic

downgrading of a suggestion invites collaborative inquiry into the suggested alternative and gives the participants a choice of either accepting or rejecting it, playing an important role in the joint decision-making process (Koester, 2002). That is, it makes the suggestion available for negotiation and opens up a group discussion on it. By downgrading their suggestions, the participants do not implicate epistemic priority as the first suggestion-maker (Heritage & Raymond, 2005) but instead they invite others' contributions to the change process, sharing the responsibility for the decision. This is because the participants have equal epistemic access to the referent situation (e.g., the reading text, instructional events), they display K+ position in terms of their English language expertise as they all orient to their content knowledge in and through talk and there are certain testing principles that the participants in the IR group adhere to in the construction of a language test, sharing a common ground (Heritage, 2013). Thus, the suggestion-makers avoid being assertive about their suggestions and leave room for the other participants' inquiry on and evaluation of the suggestion, allowing for a careful thought on the offered alternatives and a 'healthier' decision-making.

According to Stevanovic (2012), in joint decision-making settings, the proposed future action is presented in such a way that its actualization is made contingent upon the recipients' approval. That is, an orientation is displayed to the other participants' rights and obligations to be involved in the decision-making process (Heritage, 2013). As Stevanovic (2012) states, when making proposals, people "invite others to approve what has been proposed and suggest that the decisions should be reached together" (p.780). Furthermore, by downgrading their suggestions, the participants display not only other-attentiveness but also take into consideration issues of face (Brown & Levinson, 1978; Goffman, 1955) and self-other relations (Heritage, 1988). They save their negative face by avoiding imposition in such a joint decision-making task as well as their positive face so that when they are rejected, their face is not threatened (Brown & Levinson, 1978). As a result, they orient to both transactional goals by facilitating the collaborative item revision process and relational goals by avoiding infringement on the other participants' freedom of action (Koester, 2006).

Similar to the way account-giving is accomplished in phase PI, in phase SC, once a suggestion is somehow made, the participants' account-giving to (1) explain a suggestion, (2) challenge it, and (3) to display mutual agreement with and understanding of it considerably affect the decision-making process regarding the change to be made in the problematized test item. Thus, the talk-in-interaction revolves around the evaluation of the

offered suggestion by the group members, as a result of which the suggestion is either accepted or rejected. In this respect, phase SC, too, demonstrates the proactive, creative and problem-solving function of accounts (Firth, 1995).

As seen in all the extracts except for extracts 1b and 3b, the participants might provide an account for the suggestion that they make (See Table 25.3). They justify their suggestions based on evidence before relinquishing the floor and making relevant next the other participants' display of stance towards the suggestion. The findings show that the suggestion that is justified based on evidence is more likely to be accepted and put to use by the group. As presented in Table 16 in Section 4.2.1, the participants often make a thorough evaluation of an offered suggestion and provide account for it from different perspectives and dimensions, which is then met with the group's agreement and uptake. As in phase PI, they present evidence from the available shared sources such as the reading text, listening notes, and the story book as external sources of authority (Jacknick, 2013) and make reference to past instructional events (Can Daşkın, 2017; Can Daşkın & Hatipoğlu, 2019; You, 2015), the students' present state of L2 knowledge and skills as well as their level of proficiency in order to prove the appropriacy and validity of the suggestion in terms of eliminating the problem identified in the language test item.

In addition, while providing an account for a suggestion, the participants happen to orient to not only the violated testing principles that were revealed in phase PI but also other testing principles that the offered suggestions would simultaneously fulfil (See Section 5.2). Through their accounts, the participants show that their suggestion addresses not only the problem identified in the item but also other assumed testing principles that would further improve the validity and reliability of the item by, for example, being more plausible while at the same time establishing "only one correct answer" principle (extract 7b). In this way, they prove their suggestion as a valid solution for the identified problem in the item and at the same time display their situated understanding of how a test should be. Furthermore, the findings have revealed that account-giving serves functions other than justifying a suggestion based on evidence and ensuring that the problem is adequately addressed as it (1) displays a stance of certainty towards the offered suggestion (See Table 15.4 and 15.5 in Section 4.2.1), (2) establishes shared understanding in the group with respect to the suggested change, and (3) clears up a misunderstanding about the suggested change as in extract 6b.

As another interactional action contributing to the decision-making process in phase SC, a suggestion might be challenged and a stance of disagreement might be displayed with regard to it as in extracts 1b, 4b, 7b (Nissi, 2016; Wasson, 2016) (See Table 25.4). The participants are found to challenge not only the other participants' suggestions but also their own suggestions, which results in the withdrawal and self-cancellation of one's own suggestion. In other words, through self-challenges, they disagree with 'oneself' (Smithson & Diaz, 1996). In challenging an offered suggestion, the participants present evidence from the available shared sources (e.g., reading text) and make claims with reference to the violated testing principles announced in phase PI as well as other testing principles that the suggestion would further violate (See Table 17 in Section 4.2.1). As a result, they approach the suggestions critically and challenge them in terms of their failure not only in eliminating the identified problem in the item but also in fulfilling other testing principles for writing selective response items (See Section 5.2). In this way, they present the problem with the suggestion as not idiosyncratic but as going against larger disciplinary principles (Nissi, 2016; Waring, 2017).

The act of challenging an alternative by both the suggestion-maker and the other participants in response to an offered suggestion can be considered to be an important contribution to the validation process since it prompts a critical and careful thought on the suggestion in and through talk (Fulcher & Davidson, 2007) and entails information sequences where the suggestion is discussed and negotiated through accounts to resolve a work-related problem (Firth, 1995; Wasson, 2016). By means of their accounts with an orientation and reference to the test material, the participants legitimise their rejection of the suggestion and lay out the reasons for why it would not be a good idea, establishing the plausibility of their arguments (Nissi, 2016; Sharma, 2012). Consequently, such a collaborative way of evaluating and understanding the suggestions leads to the final decision with respect to the change to be made in the problematic item "in open discussion, through a process of dialogue and disagreement" (Fulcher & Davidson, 2007, p.21).

As in phase PI, the act of (self)-challenging a suggestion has sequential consequences and implications for decisions about the change to be made in the problematized test item (See "Outcome" in Table 17, Section 4.2.1). The participants (1) self-cancel their own suggestion by backing down and withdrawing it in the form of a self-initiated self-repair (extracts 1b, 4b, 7b), (2) reject an offered suggestion and then initiate a new suggestion-making sequence,

proceeding with their search for an alternative and orienting to different suggestions for change in the problematized item (extracts 1b, 4b, 7b), or (3) accept the offered suggestion by counter-challenging the disagreeing participant in a non-minimal post-expansion sequence (Stivers, 2013) and conjointly providing account and elaboration in support of the suggestion, which disproves the presented challenge and builds a consensus in the group about the appropriacy of the suggestion in terms of overcoming the problem in the test item (extracts 1b, 7b). The findings show that the challenge presented to a suggestion is often accepted and thus, the suggestion is directly rejected by the group or withdrawn by the suggestion-maker without being further argued for since there is clear evidence for its unacceptability and the issue does not require further negotiation.

However, it is seen that in those instances where the challenge is rejected and the suggestion is eventually accepted (extracts 1b, 7b), the act of challenging is framed in such a way that it prompts a group discussion on the validity of the suggestion. Rather than challenging by displaying disagreement, the participants demand some backing for their claim (Smithson & Diaz, 1996). That is, playing the devil's advocate (Leyland et al., 2016; MacDougall & Baum, 1997), they display hesitation and uncertainty with respect to the appropriacy of the offered suggestion without presenting specific evidence for their claim but inviting the group members to think about the issue they have raised before deciding on the suggestion and stimulating discussion (MacDougall & Baum, 1997). In relation to this, the issue is about the availability of the content of the newly suggested distractor in the reading text and requires a careful analysis of the context given in the text, which might be open to different interpretations (Sharma, 2012). As a result, the participants' claim is not directly rejected and neither is it accepted without further information (Smithson & Diaz, 1996). As Peercy et al. (2015) have found in their study on curricular design, the act of challenging by playing the devil's advocate creates a space for the teachers to consider different perspectives, externalize their thinking and build on one another's ideas before generating a shared and a better understanding about the offered suggestion. In Davidson and Lynch's (2002) study on the group dynamics of testcrafting, too, devil's advocate was one of the strategies the groups used in questioning each other's ideas and checking the effectiveness of the explanations offered. They point out that such strategies as playing the devil's advocate in group interaction generate a more extended discussion of alternative solutions, increasing the quality of decision-making in testcrafting.

Different from phase PI where disagreement is often managed in a way that agreement is pursued with regard to the problem in the test item and thus, the participants eventually concede the announced problem, in phase SC, the participants make no such attempt to particularly agree with the suggestion and they happen to directly agree to disagree with the suggestion after being challenged with concrete evidence. Disagreement with a suggestion in the context of a task-oriented event is found to be acceptable and unmarked and in fact, valued in the decision-making process as in Angouri's (2012) study on problem solving meeting talk. When faced with disagreement, the suggestion-makers often do not contest the group members' negative evaluation but rather confirm their contrastive position by abandoning their suggestion (Wasson, 2000), displaying agreement, or proceeding with their search for another alternative. Hence, even though sequences with disagreement are "expansion-relevant" (Schegloff, 2007), in phase SC, challenged suggestions could often be easily rejected and eliminated by the IR group without extra interactional work and sequence expansions as seen in phase PI. Consequently, different from phase PI, the participants do not show discomfort with disagreement in phase SC and contrasting positions with regard to a suggestion are not produced as a dispreferred activity (Arminen, 2005).

To illustrate, negative evaluations of the suggestions are not mitigated (Kotthoff, 1993) unlike the negative evaluations of the test items and the testers' disagreeing responses to them in phase PI where they are framed cautiously with various forms of mitigation and delaying devices. This is because in phase PI, negative evaluations are directed towards the testers as the ones who have produced the test material and hence, it is not shared in the sense that it is in phase SC where the item is revised collaboratively through group work with the feedback of all the participants including the suggestion-makers on the offered suggestions. Considering the testers' responses to the negative evaluations of their items in phase PI, becoming accountable for the announced problem in their test item, they might challenge it but usually with a form of mitigation (i.e., agreement-prefaced disagreement), orienting to disagreement with the problem as a dispreferred act in such a context (Pomerantz, 1984a; Pomerantz & Heritage, 2013; Schegloff, 2007).

On the contrary, in phase SC, no mitigating action is performed when challenging and disagreeing with a suggestion. In some instances in phase SC, even the inter-turn gaps which are normally associated with dispreferred responses (Schegloff, 2007) are not oriented to as delaying devices that project a dispreferred response to the offered suggestion. This is



because the participants including the suggestion-makers are all focused on the test material during this gap of silence carefully checking the appropriacy of the offered suggestion with their gazes at the test material. The gap can also be followed not only by the other participants' challenges but also by the suggestion-makers' own disapproval and withdrawal of the suggestion they have offered. In relation to this, the suggestion-makers' self-challenges and display of disagreement with 'oneself' (Smithson & Diaz, 1996) can take place immediately following the articulation of the suggestion without any gap of silence and without waiting for others' evaluation (e.g., extracts 4b, 7b). This further shows that they are not defensive about their suggestions and are ready to accept any challenges against them in such a task-oriented interaction. As Smithson and Diaz (1996) indicate, disagreeing with 'oneself' is a useful action for producing possible solutions to a problem and the participants' such a shift of position can take place without threatening their "sense of argumentative self because the aim of the discussion is not for participants to adhere to their positions, but rather to get a piece of work done" (p.265).

In a similar vein, in phase SC, challenging and rejecting a suggestion is not oriented to as breaking the continuity in the suggestion-making sequence (Schegloff, 2007) because phase SC quickly restarts for another round of suggestion-making as the participants continue their search for an alternative and use a preface that projects and initiates a new suggestion-making sequence. As a result, one suggestion paves the way for another suggestion in the talk-in-interaction as the participants are constantly in search of a better alternative for a change in the item. With the formulation and evaluation of each suggestion in and through talk, the participants make an attempt to produce a better one that does not have the weaknesses of the previous ones and create opportunities for immediate feedback in the IR group. When the first suggestion(s) fail in eliminating the problem in the item or raise other problems, the group realizes that the problematized part of the item requires a different change. This further proves that disagreeing with a suggestion in the context of phase SC is not treated as a dispreferred response but as an acceptable and a constructive one which the participants benefit from in formulating their suggestions for the task of item revision (Nissi, 2016). In this respect, opposing views form an inherent part of phase SC (Angouri, 2012). As Kangasharju (2002) states, "disagreement can sometimes be regarded as the responsibility of the members of an institutional committee, especially when an adjustment or correction is due" (p.1452). It is, thus, found to be a task-supporting activity that avoids poorly judged decisions in institutional settings (Angouri, 2012; Kangasharju, 2002).

The difference between phase PI and SC in the way disagreement is managed can be explained with their goal orientations in jointly accomplishing the task of reviewing language test items and revising them where necessary. In phase PI, the goal is to carefully evaluate the item and identify any potential problems in the item, making sure that it is valid, whereas in phase SC, the participants seek the best possible solution so that the item can be revised in such a way that it poses no risk for the validity of the item and that the problem is solved in the best way possible. In phase PI, challenges are presented towards the negative evaluation of the test item being reviewed while in phase SC, they constitute the negative evaluation itself with respect to an offered suggestion. As Koester (2006) explains, evaluative language is an integral part of task accomplishment in many workplace interactions. In language test IR interactions, too, the evaluation of either the item in phase PI or of the suggestions in phase SC is central for collaborative item reviewing and revising activity. In this respect, evaluations, whether positive or negative, have a transactional goal and are resolution-oriented in IR interactions (Firth, 1995; Koester, 2006). According to Lindström and Mondada (2009), in professional and institutional settings, evaluative practice functions “as a key feature for achieving the overall activity and are oriented to as well as expected as such by participants” (p.304). Therefore, the way disagreement is managed in both phases demonstrates that negative evaluations which are formulated through clear accounts and evidence either in the base FPP (phase PI) or SPP (in response to a suggestion in phase SC) are expected to be met with the group’s agreement for the accomplishment of the task at hand so that no risk is taken in such a delicate work as language test construction.

It is also seen that despite different preferences for disagreement in phases PI and SC, the act of challenging and display of an oppositional stance can change the course of action in different ways particularly when the participants agree with the disagreements and accept the challenges presented towards a claim or a suggestion. In other words, disagreements can affect the decisions made about the test items considerably and play an important role in establishing mutual understanding in the group. This means that discussing a matter at hand with multiple perspectives is important in language test IR interactions as long as concession is sought and achieved so that some subsequent action can be taken regarding the test item being reviewed. As Myers (1998) indicates, if a group discussion is dominated by a preference for agreement with assessments, it would not serve the “purpose of exploring a range of different attitudes and responses” (p.96). Therefore, depending on the local

contingencies of the task activities in institutional settings, even if the participants do disagreement as dispreferred, it is often encouraged and allowable as a way of displaying opinions while being managed against a background of consensus and shared understanding (Myers, 1998).

In this study, too, in managing the disagreements that arise, in both phases, the IR group jointly makes decisions and displays an orientation to consensus through a reversal of opinion or abandonment of position (Sharma, 2012; Wasson, 2000). They, for instance, end up agreeing with the problem (phase PI) or the suggestion (phase SC) after challenging them and then being counter-challenged by the other participants, or agreeing to disagree with an offered suggestion (phase SC) or with the announced problem (phase PI) when they are challenged and negatively evaluated by the other participants. As a result, when differences arise between them, the IR group members work towards a unity of perspectives and mutual alignment, indicating a consensus-oriented decision-making and a preference for agreement in the IR group (Sacks, 1987; Wasson, 2000). This can also be explained by the dimension of social epistemology proposed by Drew and Heritage (1992) for the analysis of institutional talk. Professionals in institutional interactions often design their talk in order to maintain a professional cautiousness and avoid taking firm positions while at the same time conveying a specific expert authority by making claims within a relevant knowledge domain (Drew & Heritage, 1992; Heritage, 1997).

The findings reveal that throughout phase SC, the participants often go beyond the use of an agreement token and display mutual agreement with and understanding of an offered suggestion or a claim made about it, whether positive or negative, through a collaborative process, 'sharing' the task of accounting for a suggestion as in phase PI (Baraldi, 2013; Firth, 1995) (See Table 25.5). In doing this, they use a variety of interactional resources such as anticipatory completion (Lerner, 1996a, 1996b, 2004; Sert, 2019; Szczepek, 2000; Van Kruiningen, 2013), additional account or evidence, candidate understanding (Antaki, 2012; Heritage, 1984b), collaborative writing activity, full or a partial modified repeat (Kim, 2002; Stivers, 2005), (re)formulation (Baraldi, 2013; Barnes, 2007; Drew, 2003; Franco & Nielsen, 2018; Heritage & Watson, 1979; Van Kruiningen, 2013), candidate explanation, choral co-production (Lerner, 2002), and reflection on the item writing process (See Table 18 in Section 4.2.1). Additionally, the findings demonstrate that the interactional act of accounting and further elaboration of the suggested change might continue in a non-minimal

post-expansion sequence (Stivers, 2013) even when an agreement is displayed with regard to the suggestion, which normatively does not project an expansion as in extracts 2b, 4b, 6b, 9b, and 10b (Schegloff, 2007).

Through such collaborative actions, the participants demonstrate their shared understanding of the suggestions and claims made about them (Sacks, 1992) and at the same time, discuss the suggested changes jointly, which facilitates dialogic coordination in the group (Baraldi, 2013) and affects the decision-making process to a considerable extent in phase SC. Thus, those interactional resources lead to a collaborative construction of arguments and a better understanding of the suggested changes in a problem-oriented context where the aim is to get the job done rather than persuade others of one's own views (Smithson & Diaz, 1996; Van Kruiningen, 2013). Furthermore, by means of such interactional resources, the participants do not only show mutual agreement and shared understanding but also claim their epistemic right to confirm a claim, indexing epistemic authority and control over the matter at hand since they all have equal epistemic access to the test material, and are in K+ position in terms of their English language expertise and testing knowledge, sharing a common ground with regard to the institutional expectations for language test construction (Heritage & Raymond, 2005; Stivers, 2005).

The suggestion-making process is seen to be such a dynamic process that it does not end when a decision is made on a change because one change in the item might require another change as in extracts 8b and 10b. In those instances, once the group decides on a major change that addresses the trouble source in the item, they are found to be in a position to proceed with the change process in the rest of the interaction by initiating a new suggestion-making sequence for a different part of the item (e.g., the subject and object of the revised sentence in extract 8b) to make the change in the item complete. Thus, they get engaged in a collaborative text production once an entry proposal for a textual change in the test item is accepted (Nissi, 2015). Such a process demonstrates that the participants should be alert to any other problems that might emerge after deciding on a particular change to be made in the test material. As a result, the change process in the revision of a problematized item in an IR interaction can be much more complex and dynamic, requiring a careful and critical eye.

In phase SC, it is also found that the interaction does not end once a decision is interactionally made in the IR group as the change needs to be noted on the exam paper to

make the phase complete. That is, the item should be revised in practice with the presence of all the participants. According to Stevanovic (2012), the act of writing down the decision demonstrates the participants' commitment to the proposed action, which brings the suggestion and decision-making sequence to the closure. At this stage, while making the change, a collaborative writing activity might take place as in extracts 1b, 8b, and 10b where the participants note down the decided change on their exam papers by inserting the accepted change into the item/text and reading aloud the revised part of the item/text as a whole through collaborative turn completions. As in extracts 4b, 5b, 6b, 7b, and 8b, the tester herself might also read aloud the modified part of the item/text or restate the kind of change that is decided while or just before noting it down on her paper. These resources of collaborative writing and reading aloud enable a final check for the accepted suggestion and ensure shared understanding in the group with respect to the change made in the item, enacting the jointness of the activity. That is, the final entry in the text is marked as the joint view of the IR group and as the final decision through the interplay between verbal, embodied and material resources (Nissi, 2015).

In phase SC, the interactional actions discussed so far construct the jointness in decision-making since they demonstrate the three components required to arrive at a joint decision, namely access, agreement, and commitment (Stevanovic, 2012). According to Stevanovic (2012), joint decisions are constituted when the recipient “establishes *access* to the content of the proposal; expresses *agreement* with the proposer's views; and displays *commitment* to the proposed future action” (p.781). In this study, the actions of seeking and carrying out collaborative inquiry about the offered suggestion with an orientation to the test material, reformulating it through collaborative turn sequences, giving account for it, challenging it and displaying mutual understanding show the participants' access to the content of the suggestion, as a result of which they can decide whether to reject or accept it based on common knowledge. When they display agreement with an offered suggestion through a collaborative process after evaluating and understanding the suggestion, they make their positive stance with respect to it clear. Upon their approval of the suggestion, the tester then writes down the decision on her exam paper often through a collaborative writing activity or reading aloud, which marks the group's commitment to the proposed action in turns at talk as in Stevanovic's (2012) study.

This study has also shown that the proposal can be abandoned before the three components are completed as one of the outcomes of the decision-making process (i.e., non-decision) (Stevanovic, 2012). This emerges between the components of access and agreement when the participants establish access to the content of the proposal but display disagreement rather than agreement by challenging and rejecting the suggestion as a result of which the suggestion-makers often refuse to pursue the same proposal. Different from Stevanovic's (2012) finding where the proposal is abandoned due to the recipients' lack of action and from Nissi's (2016) study where the proposal is rejected implicitly through a conditional construction, in this study, an offered suggestion is more explicitly and directly rejected and the participants make their negative stances towards the suggestion more apparent through accounts. As explained above, this could be due to the participants' orientation to disagreement with a suggestion not as a dispreferred response and their responsibilities, rights and obligations (Heritage, 2013) to give immediate feedback and make their stance towards a suggestion clear so that they can proceed with their search for the best possible alternative to revise the problematized test item in such a task-oriented problem-solving context.

#### **5.1.2.3. Reviewing the Change (RC)**

RC as the third and final phase of language test IR interactions is about the participants' review of the change that is decided and made in phase SC to overcome the problem identified in the item in phase PI. In this phase, the IR group checks and finds out whether the revised item complies with the testing principles for language test construction, addressing the violated testing principle and causing no additional problems that could go unnoticed with the revision of the item. Such a stage is vital considering the significance of the product that is produced in this task-oriented IR interaction. This phase is found to be operationalized through a particular recurring interactional pattern as illustrated in a model in Figure 5 (See Section 4.3.1), which is often repeated several times in the extracts.

As Table 26 lists, there are certain recurring interactional actions that help the group to interactionally review and re-evaluate the item after it has undergone a change in phase SC and that, in fact, facilitate the item change reviewing process. As a result of such a process, the IR group eventually confirms and finalizes the revision of the item, ending the IR interaction. The sequential position of the interactional actions presented in Table 26

changes. The acts of seeking confirmation and challenging the change occur alternatively as two possible sequence-initiating actions in phase RC. These actions may be followed by the act of giving account for the change in the item, which is the third action in Table 26. However, the sequence-initiating act of challenging the change more often makes relevant next the act of suggesting a new change where phase SC is reinstated (Table 26.4). As for the act of displaying mutual agreement and understanding as the last item in Table 26, it takes place throughout phase RC in response to a claim made about the revised item. In this section, those interactional actions are discussed respectively as given in Table 26.

**Table 26. Recurring Interactional Actions in Phase RC**

| No | Interactional Actions                            | Extract No      |
|----|--------------------------------------------------|-----------------|
| 1  | Seeking and checking confirmation for the change | 1c, 3c, 4c      |
| 2  | Challenging the change                           | 1c, 2c, 4c, 10c |
| 3  | Giving account for the change                    | 1c, 3c, 4c      |
| 4  | Suggesting a new change (Phase SC reinstated)    | 1c, 2c, 4c, 10c |
| 5  | Displaying mutual agreement and understanding    | All extracts    |

There are two possible sequence-initiating actions that lead to the emergence of phase RC and enable the group to reconsider the revised item; (1) confirmation requests (extracts 1c, 3c, 4c) and/or (2) challenging of the change (extracts 1c, 2c, 4c, 10c). To begin with the former, it is seen that in this phase, the participants seek and check the group's final confirmation for the revised item often using a tag question that mobilizes support for their claim regarding the change made in the item (Heritage, 2012). Their confirmation requests function as a devil's advocate as they invite the group to re-evaluate the change and hence, lead to a collaborative process of reviewing the change, generating further discussion about its appropriacy (Davidson & Lynch, 2002; Leyland et al., 2016; MacDougall & Baum, 1997; Peercy et al., 2015). In this way, although a decision addressing the problem was already made and confirmed by the group in phase SC, the participants once again (1) make sure that the change fulfils not only the testing principle that is claimed to be violated in the first phase of the IR interaction but also other testing principles for language test construction, causing no additional problems (See Section 5.2) and (2) establish shared or correct understanding with regard to the revised item before ending the reviewing interaction for this particular item.

Another sequence-initiating action that shapes the talk-in-interaction in phase RC is challenging the change as seen in extracts 1c, 2c, 4c, and 10c (See Table 26.2). While reviewing the change in the test item, the participants recognize a problem and challenge the revised item with reference to testing principles for item writing (See Section 5.2), which grounds their claims in disciplinary or professional experience and expertise (Nissi, 2016; Waring, 2017). The participants identify a new problem that emerges in the item as a result of its revision in phase SC and that violates a testing principle that is different from the one raised in phase PI (See Table 20 in Section 4.3.1). For instance, through the act of challenging the change, the focus of the interaction is shifted from the content of the option as the main concern onto its format or grammaticality.

The participants' confirmation requests and act of challenging the change have sequential consequences and implications for decisions about the revised item. The group either confirms the change or agrees with the challenge presented towards the change, reinstating phase SC to solve the emerging problem in the revised item. In the case of the former, the item is confirmed often following a process of account-giving in support of it (extracts 1c, 3c, 4c) (See Table 26.3). As in the previous phases, the findings show that the participants provide account for the change in the item collaboratively from different perspectives and confirm the revised item based on evidence. In other words, through account-giving, they make sure that the problem is solved and the testing principle in question is put to use without posing any problems. Furthermore, while giving account, the participants bring their professional expertise to the interaction and show their subject knowledge as language teachers (e.g., knowledge of conjunctions in extract 3c) or present concrete evidence from the shared sources such as the reading text as external sources of authority (Jacknick, 2013) (extract 4c). In addition, as seen in extract 4c, even though the participants' display of agreement with regard to the revised item does not project an expansion (Schegloff, 2007), the act of account-giving for the already decided and noted change in the item might continue in a non-minimal post-expansion sequence (Stivers, 2013) through a collaborative process due to the goal-oriented character of the interaction (Van Kruiningen, 2013). Consequently, such a collaborative and comprehensive process of account-giving leads to the acknowledgement and approval of the change in open discussion as well as the establishment of shared understanding about the revised item.



In the case of the latter when the revised item is challenged and disconfirmed by the group, the item undergoes another change through a new suggestion and decision-making process (See Table 26.4), following the same structural organization that has appeared in phase SC (extracts 2c, 4c, 10c). It is only in extract 1c where the presented challenge is rejected and the change is then confirmed through collaborative account-giving because the participant DE challenges the change in the form of a devil's advocate (Davidson & Lynch, 2002; Leyland et al., 2016; MacDougall & Baum, 1997; Peercy et al., 2015) with no specific evidence by displaying hesitation accompanied by her thinking face and raises an issue for the group members to all consider before approving the item. In other words, DE challenges the suggestion by demanding some backing or justification for her claim rather than displaying disagreement (Smithson & Diaz, 1996). The issue is also about the availability of the content of the revised option in the reading text, which requires a careful analysis of the context given in the text and the discussion of possible interpretations (Sharma, 2012).

In all the other extracts, the participants accept the challenge that has made the problem in the revised item clearly visible (e.g., ungrammaticality of the sentence) and move towards solving the emerging problem through a suggestion and decision making process in the rest of the interaction in phase RC (See "Outcome" in Table 20, Section 4.3.1). As a result, the group decides on a new change to overcome the problem that has emerged after the revision of the item and the item is eventually finalized. This indicates that language test IR interactions require continuous reviewing and editing since a major change might lead to other changes to make the revision of the item complete. This could happen either within the phase of SC before the decided change is fully acknowledged and noted on the exam paper as seen in extracts 8b and 10b or in a separate phase as in the extracts in question (i.e., 1c, 2c, 4c, 10c) where the participants review the already noted change in terms of different testing principles before proceeding with the other items in the test. Consequently, the findings show that the change process in the revision of a problematized item in an IR interaction can be much more complex and dynamic, requiring a careful and critical eye.

Considering the challenges that are accepted, it is seen that the matter at hand is displayed and regarded as assumed knowledge with respect to the testing principles adopted for language test construction within the institutional context in question. This is because the raised issue is not further questioned and discussed as the group directly orients to a search for a change to overcome the problem, which signals their already acceptance of the

presented challenge. As in phase SC, the act of challenging is, thus, treated as a preferred act and is often accepted directly without extra interactional work. This is because the goal in phase RC is to do a final check for the item and make sure that the revision of the item is complete and the item is in line with the testing principles, causing no further problems. The act of challenging foresees future problems that might possibly arise once the test is administered and is, hence favoured by the group members. In relation to this, since the revision of the test is based on joint work and decision in the group, the participants can exchange their views in open dialogue without necessarily addressing their negative evaluations towards a particular person as in phase PI where the identified problem is addressed to the tester and thus, is framed cautiously.

Finally, as in the previous phases, the findings show that throughout phase RC, the participants often go beyond the use of an agreement or a compliance token and display mutual agreement and understanding through certain interactional resources (See Table 26.5). This is important in establishing shared understanding and making the group's stance clearer so that the change can be overtly confirmed and a decision can be made jointly in such a shared task before ending the IR interaction. In doing this, the participants use a variety of interactional resources such as additional account or evidence, repetitional response (Heritage & Raymond, 2012), modified repeat (Kim, 2002; Stivers, 2005), (re)formulation (Baraldi, 2013; Barnes, 2007; Drew, 2003; Franco & Nielsen, 2018; Heritage & Watson, 1979; Nissi, 2015; Van Kruiningen, 2013), candidate explanation, positive assessment, and anticipatory completion (Lerner, 1996a, 1996b, 2004; Sert, 2019; Szczepek, 2000; Van Kruiningen, 2013) (See Table 22 in Section 4.3.1). As a result, these interactional resources overtly confirm the revised item or the issue raised about it and eventually lead to the final decision through a collaborative construction of arguments (Smithson & Diaz, 1996; Van Kruiningen, 2013) and dialogic coordination (Baraldi, 2013), ending the IR interaction. Furthermore, by means of such interactional resources, the participants do not only demonstrate mutual agreement and shared understanding (Sacks, 1992) but also claim their epistemic right to confirm the revised item or the evaluations and suggestions made about it, indexing epistemic authority and control over the matter at hand (Heritage & Raymond, 2005; Stivers, 2005). This is because they all have equal epistemic access to the test material, and display their K+ position in terms of their English language expertise and testing knowledge, sharing a common ground with regard to the institutional expectations for language test construction (Heritage, 2013).

## **5.2. Testing Principles Oriented to Throughout the Phases of Language Test IR Interactions**

Throughout the three phases of language test IR interactions examined in this study, the participants orient to a wide range of testing principles described in the literature for item writing in general and for selected response (i.e., multiple-choice, matching) or limited production (i.e., fill-in-the-blanks) item formats in particular (e.g., Carr, 2011; Fulcher, 2010; Haladyna & Rodriguez, 2013; Harris, 1969; Hatipoğlu, 2009; Heaton, 1990; Kehoe, 1995). This shows that there are certain testing principles that the participants seek in the construction of syllabus-based language tests in their institution. By orienting to those testing principles, they base their claims on reasonable and professional grounds. According to Waring (2017), by invoking larger principles, professionals ground their claims in disciplinary experience and expertise, ensuring understanding of the matter at hand. In this study, the participants make reference to testing principles most notably while (1) announcing and explaining the problem in the test item in phase PI, (2) providing account and evidence for their suggestions as well as challenging them in phase SC, and while (3) seeking confirmation and account-giving to confirm the change as well as challenging the change in the revised item in phase RC. In those instances, the oriented testing principles are claimed to be either violated or fulfilled by the test item and the participants move towards solving the problem if the item is found to violate a testing principle.

As shown in Table 27, one of the major violated testing principles that the participants make reference to while giving account for the problem in the language test item in phase PI is the possibility of more than one correct answer (extracts 3a, 4a, 6a, 7a, 9a, 11a) (Carr, 2011; Haladyna & Rodriguez, 2013). This is followed by “test what you teach” principle (extracts 2a, 8a, 12a) (i.e., the item should represent a specific content area covered in the course) in which case the participants identify any mismatches between the test and the instructional events or the syllabus/course content, creating the interplay between teaching and testing and establishing temporality and continuity in L2 educational activities (Can Daşkın & Hatipoğlu, 2019). They also orient to the violation of testing principles concerning (1) the plausibility of options in selected response items (extract 1a), (2) the simplicity of vocabulary in a reading text (extract 8a) so that “reading comprehension does not interfere with testing the content intended” (Haladyna, 2004, p.99), (3) the appropriacy of the language used in the item (extract 5a), (4) the use of ambiguous terms in the item (extract

9a), (5) the importance of the content that is tested (extract 10a), (6) the difficulty level of the item considering the students' language skills and knowledge (extract 12a), and (7) the independency of the item (extract 12a). This demonstrates that the reviewers analyze and evaluate the language test items from various dimensions, drawing on not only their teaching experience but also their testing and language expertise. That is, they focus not only on the content of the item (i.e., principles of “test what you teach” and “avoidance of trivial content”) and their students' L2 skills and knowledge in line with the teaching and learning context (i.e., principles of “simplicity of vocabulary” and “level of difficulty”) as their main goal in the IR interaction but also on language-related issues (i.e., principles of “appropriacy of language” and “avoidance of ambiguity”), key verification (i.e., principle of “only one correct answer”) as well as more testing-related issues (i.e., principles of “plausibility” and “independency”).

**Table 27. Testing Principles Oriented in Language Test IR Interactions**

| No | Testing Principle            | Description                                                                                                                    | Extract No                                      |
|----|------------------------------|--------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------|
| 1  | Only one correct answer      | Each item should have only one correct answer. (Carr, 2011; Haladyna & Rodriguez, 2013; Hatipoğlu, 2009; Heaton, 1990)         | 3a, 4a, 6a, 7a, 9a, 11a, 4b, 6b, 7b, 9b, 3c, 4c |
| 2  | Plausibility                 | Options should be reasonably attractive and plausible. (Carr, 2011; Haladyna & Rodriguez, 2013; Hatipoğlu, 2009; Heaton, 1990) | 1a, 1b, 4b, 7b, 4c, 10c                         |
| 3  | Test what you teach          | The item should represent a specific content area covered in the course. (Carr, 2011; Haladyna & Rodriguez, 2013)              | 2a, 8a, 12a, 2b, 1c                             |
| 4  | Level of difficulty          | The item should be at a level appropriate to the proficiency level of test-takers. (Heaton, 1990)                              | 12a, 5b, 4c                                     |
| 5  | Appropriacy of language      | The language used in the item should sound as natural as possible. (Harris, 1969)                                              | 5a, 5b                                          |
| 6  | Simplicity of vocabulary     | The vocabulary in a reading text should be kept simple. (Haladyna, 2004; Harris, 1969)                                         | 8a, 7b                                          |
| 7  | Avoidance of trivial content | Each item should be based on important content. Trivial content should be avoided. (Haladyna & Rodriguez, 2013)                | 10a, 10b                                        |
| 8  | Avoidance of “no answer”     | The item should be checked for an absence of correct answer. (Hatipoğlu, 2009)                                                 | 1b, 1c                                          |

**Table 27 (continued)**

|    |                                      |                                                                                                                                                                                                     |        |
|----|--------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| 9  | Odd man out                          | Options should be similar in structure. (Carr, 2011; Fulcher, 2010; Hatipoğlu, 2009; Haladyna & Rodriguez, 2013)                                                                                    | 4b, 4c |
|    |                                      | The length of options should be kept about equal. (Haladyna & Rodriguez, 2013)                                                                                                                      |        |
| 10 | Window dressing                      | Excessive verbiage or irrelevant clues should be avoided in the item. (Haladyna, 2004; Hatipoğlu, 2009; Heaton, 1990)                                                                               | 6b     |
| 11 | Avoidance of ambiguity               | Ambiguous terms and structures should not be used in the item. (Carr, 2011; Hatipoğlu, 2009)                                                                                                        | 9a     |
| 12 | Independency                         | The content of individual items should be independent of each other. (Haladyna & Rodriguez, 2013)                                                                                                   | 12a    |
| 13 | Sufficiency of context               | There should be sufficient context to convey the intent of the question to the students in a fill-in item. (Brown, 1996)                                                                            | 6b     |
| 14 | Identification of acceptable answers | A list of acceptable answers should be provided for productive response items. (Brown, 1996)                                                                                                        | 6b     |
| 15 | Grammaticality                       | Options should form a grammatical sentence when placed in the stem. (Hatipoğlu, 2009; Heaton, 1990)                                                                                                 | 2c     |
| 16 | Text coherence                       | In a reading comprehension test, the text should allow test-takers to understand the relation between parts of a text through both lexical devices and grammatical cohesive devices. (Heaton, 1990) | 8b     |

\*a- phase PI, b- phase SC, c- phase RC

The reviewers' orientation to a variety of testing principles in phase PI shows that they go beyond the sub-processes of content review, key check and editorial review proposed by Fulcher and Davidson (2007) and Fulcher (2010) for item reviewing. The consideration of all these issues collectively in and through interaction helps teachers identify any construct-irrelevant factors that might affect test performance and cause a threat to the validity of the item (Fulcher, 2010; Haladyna & Rodriguez, 2013). Furthermore, it is seen that the identification of one testing principle that is violated leads to the violation of the other, resulting in overlapping testing principles. To illustrate, in extract 9a, the use of an ambiguous term (i.e., *can*) in one of the distractors (i.e., principle of "avoidance of ambiguity") causes different interpretations of the distractor in the IR group and makes it

likely to be selected as a correct answer (i.e., principle of “only one correct answer”), affecting the key of the item. The language test IR interactions, thus, constitute a highly interactive multicomponent item review process (Welch, 2006).

In phase SC, through their acts of account-giving for and challenging a suggestion, the participants address not only the testing principle that is claimed to be violated in phase PI but also other testing principles for item writing. In account-giving, they refer to the testing principles that their suggestions would fulfil including the one the item was found to violate. By addressing the violated testing principles that were revealed in phase PI, the participants prove their suggestion as a valid solution for the identified problem in the item, which makes the suggestion more likely to be accepted by the IR group (extracts 2b, 4b, 5b, 6b, 7b, 9b, 10b). To illustrate, in extracts 4b and 9b, the suggestions offered for the possible correct distractor in a reading comprehension test are justified in terms of their unavailability in the reading text and thus, the violated testing principle that each item should have only one correct answer (Carr, 2011; Haladyna & Rodriguez, 2013; Hatipoğlu, 2009; Heaton, 1990). Similarly, in extract 2b, the reviewers address the violated testing principle that the item should represent a specific content area covered in the course (i.e., test what you teach) (Carr, 2011; Haladyna & Rodriguez, 2013) by explaining that the suggested verb “*change*” appeared in the course and the students are more familiar with it than the original verb “*cover*” given in the item to be transformed into its past tense form.

Additionally, by orienting to different testing principles that their suggestions would simultaneously fulfil besides the violated testing principle, the participants at the same time display their situated understanding of how a test should be and show that their suggestion would further improve the validity and reliability of the item (extracts 5b, 6b, 7b, 8b). For instance, in extract 7b, one of the suggestions (i.e., *don't have to*) for the problematized distractor option (i.e., *should*) is claimed to make the option not only incorrect as the main concern but also more plausible for the target group of students. Likewise, in extract 5b, in their accounts given for the replacement of “*before*” in the target sentence with the suggested expression “*the early years of their education*”, the participants address not only the violated testing principle concerning the appropriacy of the language used in the item but also the testing principle about the difficulty level of the item by explaining that the new expression would make the item relatively more difficult for the intended group of test-takers. In addition, in extract 6b, while explaining her suggestion (i.e., including the other possible

correct answer “*physical reactions*” in the item as a continuation of the available sentence) as a valid solution for the violated testing principle of “only one correct answer” in a fill-in-the-blank item, the participant SA at the same time refers to the testing principle that there should be “sufficient context to convey the intent of the question to the students” in a fill-in item (Brown, 1996, p.58). This would, in turn, ensure that the required response is concise (Brown, 1996) and there is only one correct answer (Carr, 2011; Haladyna & Rodriguez, 2013; Hatipoğlu, 2009; Heaton, 1990).

With respect to the act of challenging a suggestion in phase SC, the analysis has revealed that while some suggestions are challenged with reference to the violated testing principle announced in phase PI, others are found to violate a different testing principle in an attempt to solve the problem in the item (extracts 1b, 4b, 7b). This often results in the rejection of the suggestion by the IR group. In extracts 4b and 7b, for instance, the identified problem in the test items is related to the possibility of more than one correct answer and thus, the violation of “only one correct answer” principle. However, when challenging the suggestions offered to solve this problem, an orientation is shown not only to “only one correct answer” principle but also other testing principles in relation to the plausibility (extracts 4b, 7b) and the structure of the suggested distractors (i.e., “odd man out” principle) (extract 4b). In the case of the latter when the suggestion is challenged in terms of its structure, the participants point out that it would turn out to be the only negative option in the item, violating the testing principle that “options should be similar in structure” (Fulcher, 2010, p.172) and being noticed as “the odd man” (Carr, 2011; Hatipoğlu, 2009).

Furthermore, in extract 1b, the item asks for an advantage *not* mentioned in the text (i.e., *Which of the following is NOT one of the advantages of the Kalenjini?*) and the problem identified in the item is related to the implausibility of the correct option as it was found to be too irrelevant to the content of the text and thus, too obvious to select as a “not stated” option. However, in an attempt to make the correct option more plausible, the participants make sure that it is not so close to the content of the reading text that it ends up being expressed as an advantage in the text and thus, being an incorrect option. Therefore, they challenge the offered suggestions with reference to not only their plausibility as the main concern but also their availability in the text and thus, the violation of “avoidance of no answer” principle (Hatipoğlu, 2009) because “no answer” in an item would cause more serious problems than an implausible distractor. As a result, the participants disprove the

offered suggestions by showing that they would either retain the same problem in the item or cause additional problems that could go unnoticed with the revision of the item.

Overall, as in phase PI, the participants evaluate their suggestions from various dimensions in phase SC, drawing on not only their teaching experience but also their testing and language expertise. However, different from phase PI, most of the testing principles oriented to in phase SC are related to language test construction, requiring testing expertise. They include issues such as the plausibility of options (extracts 1b, 4b, 7b), odd man out (extract 4b), window dressing (extract 6b), sufficiency of context (extract 6b), and text coherence (extract 8b). Orientation to a greater number of testing-related issues could be due to the relatively more equal and active involvement of both testers and reviewers in the item revision process in phase SC as compared to the announcement of problem in phase PI which belongs to the reviewers (See Section 5.3). Thus, in such a shared task of revising the problematized test items, the participants evaluate their suggestions not only in terms of solving the problem addressed in phase PI but also in terms of their adherence to item writing guidelines (Haladyna & Rodriguez, 2013). Through such a qualitative item reviewing and revising procedure in and through interaction, the participants provide important sources of validity evidence for their suggestions, which have sequential consequences and implications for decisions about the changes to be made in the item (Downing, 2006; Weir, 2005; Wolf et al., 2008). As a result, with an orientation to testing principles in phase SC, it is ensured that (1) the suggested change is a valid solution for the identified problem in the item, (2) it does not cause additional problems, (3) the validity and reliability of the item is further improved with the revision of the item in terms of the fulfilment of different testing principles, and (4) shared understanding is established about the change as the participants display a situated understanding of how a test should be.

As for phase RC in language test IR interactions, it is seen that while seeking confirmation and account-giving for the change in the revised item (extracts 1c, 3c, 4c) as well as making a final comment about it as a sequence-closing action (extracts 2c, 4c, 10c), the participants orient to mainly the violated testing principle revealed in phase PI but they are also found to consider other testing principles for language test construction as in phase SC. To illustrate, in extracts 3c and 4c, through their confirmation requests often in the form of a tag question that mobilizes support (Heritage, 2012) and functions as a devil's advocate (Davidson & Lynch, 2002; Leyland et al., 2016; MacDougall & Baum, 1997; Peercy et al., 2015), the



participants invite the group to re-evaluate the revised distractors in terms of their violation of “only one correct answer” principle and thus, to check whether the problem is adequately addressed. Similarly, in extracts 2c and 10c, when making a final comment about the item just prior to closing the RC sequence, the participants refer to the main issue addressed in phase PI (i.e., principle of “test what you teach” in extract 2c; principle of “avoidance of trivial content” in extract 10c) in order to establish common ground on the underlying reason for the revision of the item. They make sure that the problem is solved and is clearly understood by the group just at the point when the participants shift their focus to a different issue while reviewing the revised item.

However, in extract 1c, playing the devil’s advocate, the participant DE checks the other participants’ final confirmation with respect to the availability of the revised correct option in the reading text since the item asks for an advantage *not* mentioned in the text. In this way, she displays an orientation not to the plausibility of the option as the main concern but to the principle of “avoidance of no answer”. Likewise, in extract 4c, the reviewer DU goes beyond the main issue (i.e., principle of “only one correct answer”) in her final comment about the item and remarks on the difficulty level of the revised item as well as the plausibility of the new distractor as the testing principles that are further put to use with the revision of the item.

On the other hand, when challenging the revised item in phase RC, the participants refer to the violation of testing principles that are different from the one raised in phase PI because the item was already revised in accordance with the violated testing principle in phase SC (extracts 1c, 2c, 4c, 10c). Nevertheless, what leads to the act of challenging the change is that in some instances in phase RC, the revised item turns out to violate other testing principles since with the revision of the item that addresses the major problem, some minor problems emerge in the test item, which mainly involves editorial review. To exemplify, in extract 2c, the focus of the interaction is shifted from “test what you teach” principle as the main concern onto the ungrammaticality of the target sentence where the verb has undergone a change. As for extract 4c, through the act of challenging the change, an orientation is displayed to the violation of a testing principle which states that “options should be similar in structure” (Fulcher, 2010, p.172) and more specifically that the length of options should be kept about equal (Haladyna & Rodriguez, 2013). As a result, the focus of the interaction is shifted from the content of the option and “only one correct answer” principle as the main

concern onto its format. Finally, in extract 10c, the focus of the interaction is shifted from the testing principle concerning the importance of the content that is tested in the item (i.e., principle of “avoidance of trivial content”) onto the testing principle about the plausibility of the alternative character names given in the box as distractors.

Overall, similar to those in phase SC, in phase RC, an orientation is shown to a number of testing principles before finalizing the item and ending the IR interaction. The participants do a final check of the revised item for either the fulfilment or the violation of testing principles. They bring forward similar issues discussed earlier in the previous phases. It is seen that even if the testing principle in question was already ensured to be fulfilled in phase SC, playing the devil’s advocate (Davidson & Lynch, 2002; Leyland et al., 2016; MacDougall & Baum, 1997; Peercy et al., 2015), they orient to the same principle once again in phase RC, generating further talk about it (e.g., “avoidance of no answer” principle in extract 1c; “only one correct answer” principle in extract 4c). However, an opportunity is also created to identify and solve some minor problems that arise when the item undergoes a change in phase SC (e.g., an ungrammatical sentence in extract 2c, an option that is longer than others in extract 4c, implausible options in a fill-in-the-blanks item in extract 10c), which illustrates the importance of launching phase RC for final check of the item.

Consequently, before finalizing the item and ending the IR interaction, an orientation to testing principles in phase RC helps solicit a final confirmation for the item based on a solid ground and recognize and deal with other problems that could possibly emerge as a result of revising the item (extracts 2c, 4c, 10c). It ensures that the change made in the item (1) has solved the problem (extracts 2c, 3c, 4c, 10c), (2) causes no additional problems that could go unnoticed (extract 1c, 2c, 4c, 10c), and/or (3) makes further contributions to the validity of the item with the establishment of different testing principles (extract 4c). Eventually, shared understanding is developed in the IR group with regard to the final version of the test item and the joint work of item reviewing and revising is completed successfully. According to Waring (2017), as members of the same community of practice, invoking larger disciplinary principles opens up a space for “exploring principled understandings, and socializing the teacher into important disciplinary and pedagogical conduct and conceptualizations” (p.30). As a result, the teachers collaboratively arrive at a common understanding about the fulfilment or violation of testing principles in the reviewed test items.

### **5.3. Institutional Role Distribution of Participants in the Phases of Language Test IR Interactions**

The findings show that throughout the phases of language test IR interactions, an orientation is displayed to the institutional role distribution of the participants depending on the sub-goal of each phase. The participants' institutional roles and responsibilities are demonstrated as being relevant in and through the moment-to-moment talk and accomplished situationally in ways specific to the IR activity (Drew & Heritage, 1992; Sidnell, 2013; Zimmerman & Boden, 1991). In phase PI, the roles that are made relevant in and through talk are more distinctive because the product being reviewed is 'owned' by the tester and it is the place where the teachers have the rights to evaluate and give their feedback on the test item as their main goal in the IR interaction (Heritage, 2013). However, this role distribution disappears slightly in the following phases since the group reproduces the item collaboratively as the product of shared work and their roles and responsibilities are not as stable as they are in phase PI. In other words, a new item writing process where they are all involved takes place. As a result, the social structure of the IR practice is transformed in phases SC and RC, going beyond giving feedback on the item to its co-construction through a change process. The initial task turns into one of a group work where the reviewers act as testers and the tester acts as a reviewer by offering suggestions and evaluating them. In this respect, they orient to the understood goal of doing the SC or RC task jointly as sequentially ordered action components of the IR practice (Ford, 1999).

As Table 28 shows, due to the goal of phase PI, an orientation is displayed to certain tester and reviewer roles and responsibilities that are not seen in the other phases. As the ones invited to the IR session, the reviewers have the primary authority and rights to evaluate the test items (Heritage, 2013). Therefore, they problematize the test items by giving account and evidence through some interactional actions such as carrying out preliminary inquiry about the problem. They also orient to their roles as 'problematizers' of test items in the following two phases by providing a further clarification of the problem they have announced in phase PI. This reminds the IR group of the main issue and establishes common ground on the underlying reason for the revision of the item. Moreover, the reviewers imply epistemic authority by announcing a problem in the item and addressing it to the tester with a shift of gaze to her. In this way, they project an unknowing status for her since they raise an issue about the item that might have gone unnoticed by the tester or that is about the

teaching context which the tester has a relative lack of epistemic access to. This is the reason why the reviewers perform mitigating actions while announcing and explaining the problem in the test item.

On the other hand, the testers display their role and goal specific to phase PI by (1) announcing the expected correct answer or (dis)confirming the reviewers' answers to the test items, (2) pre-emptively giving an elaborative account for the expected correct answer to provide insight into the projected problem, (3) challenging the problem where necessary by giving counter-evidence and making counter-claims and (4) being alert to any problems that might be addressed in the IR interactions with respect to language test items, which is most evident in their candidate evaluations of the item as a result of their anticipation of the projected problem (extracts 1a, 5a). The reason for the emergence of these interactional actions that belong to the testers in phase PI is that they are made accountable by the reviewers' actions concerning their product. Those interactional actions also signal the testers' epistemic right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item (Type 1 knowledge, Pomerantz, 1980) as the writer and thus, the 'owner' of the item. Therefore, they position the testers as already knowing (K+) recipients and ascribe them an epistemic authority (Heritage, 2012, 2013). As a result, it is seen that in phase PI, the interactional actions exhibit an orientation to tester and reviewer roles and responsibilities that are peculiar to phase PI.

**Table 28. Tester and Reviewer Roles and Responsibilities Made Relevant in Language Test IR Interactions**

| Phases        | Tester Roles and Responsibilities                                                                  | Extract No                       |
|---------------|----------------------------------------------------------------------------------------------------|----------------------------------|
| Phase PI      | Announcing the expected correct answer or (dis)confirming the reviewers' answers to the test items | 3a, 4a, 5a, 6a, 7a, 9a, 10a, 11a |
|               | Providing an elaborative account for the expected correct answer                                   | 1a, 10a, 11a                     |
|               | Challenging the announced problem                                                                  | 3a, 4a, 9a, 10a, 11a, 12a        |
|               | Being alert to upcoming problems in the item                                                       | 1a, 5a, 10a                      |
| Phases PI, SC | Reflecting on the item writing process                                                             | 5a, 6a, 9a, 7b                   |

**Table 28 (continued)**

| Phases SC,<br>RC     | Noting down the change on the test material                                        | All extracts in<br>phase SC,<br>2c, 4c, 10c                      |
|----------------------|------------------------------------------------------------------------------------|------------------------------------------------------------------|
| Phases PI,<br>SC, RC | Ensuring correct and shared understanding of the<br>problem or the accepted change | 1a, 3a, 8a, 2b, 4b,<br>5b, 6b, 7b, 8b, 10b,<br>1c, 2c, 4c        |
| <b>Phases</b>        | <b>Reviewer Roles and Responsibilities</b>                                         | <b>Extract No</b>                                                |
| Phase PI             | Problematizing the item through account-giving                                     | All extracts in<br>phase PI                                      |
| Phases PI,<br>SC     | Reflecting on the test-taking process                                              | 3a, 10a, 11a, 6b                                                 |
| Phases SC,<br>RC     | Providing further clarification of the problem<br>announced in phase PI            | 1b, 5b<br>2c, 10c                                                |
| Phases PI,<br>SC, RC | Framing claims from students' and teaching<br>perspectives                         | 2a, 4a, 6a, 8a, 9a,<br>11a, 12a, 2b, 7b,<br>10b, 1c, 2c, 4c, 10c |
|                      | Reading the story before the reviewing of the story<br>exam                        | 10a, 10b, 10c                                                    |

\*a- phase PI, b- phase SC, c- phase RC

The findings further reveal participant roles and responsibilities that are not specific to phase PI but that still display a clear distinction between the tester and reviewer roles in language test IR interactions. As Table 28 illustrates, whereas the reviewers reflect on the test-taking process while explaining how the problem has emerged in the item, the testers reflect on the item writing process during their act of account-giving in phases PI and SC although these two interactional resources are more commonly employed in phase PI due to the act of account-giving as the central action in this phase. In this way, the reviewers orient to their responsibility to take the test as a student would before giving feedback to the tester in the IR session (Alderson et al., 1995; Milanovic, 2002) while the testers once again signal their right, responsibility and obligation (Heritage, 2013) to know and have first-hand access to the test item as the writer of the item since the act of problematizing the item makes them accountable to the reviewers. Additionally, in phases SC and RC, the testers' non-verbal behaviour of noting down the decided change on the test material orients to their responsibility and obligation to finalize the test and make it ready for administration. This action is peculiar to phase SC and when phase SC is reinstated in phase RC due to the change process that the problematized item undergoes as the ultimate goal of this phase.

The remaining interactional actions displaying an orientation to the institutional role distribution of the participants are seen in all the three phases, which further indicates that the participants have specific roles and responsibilities in the IR interactions regardless of the sub-goal of each phase. Firstly, the findings show that the testers often ensure correct and shared understanding of the problem in phase PI or of the accepted change in phases SC and RC. In phase PI, the testers' confirmation requests in the form of a candidate explanation for (extract 3a) or a candidate understanding (Antaki, 2012; Heritage, 1984b) of the problem (extracts 1a, 8a) to clarify the matter at hand orient to their role as the writer of the item who is accountable for the test. Similarly, in phase SC, they do this for a correct and mutual understanding of the accepted change this time by (1) offering a candidate understanding (Antaki, 2012; Heritage, 1984b) of the suggestion in question (extracts 2b, 6b), (2) eliciting its repetition and dictation through a DIU (extract 10b) (Nissi, 2015) or (3) reading aloud the modified part of the item/text (extracts 4b, 5b, 7b, 8b) or reannouncing the kind of change that is decided (extract 6b). As for phase RC, the testers check and ensure correct understanding of the change that is either already accepted by the group in phase SC or newly produced in phase RC by offering a candidate understanding (Antaki, 2012; Heritage, 1984b) of the change (extract 1c) or providing a (re)formulation of the prior suggestion turn (Baraldi, 2013; Barnes, 2007; Drew, 2003; Franco & Nielsen, 2018; Heritage & Watson, 1979; Nissi, 2015) (extracts 2c, 4c). In these ways, in phases SC and RC, the participants announce the final entry in the text by recycling the preceding entry proposal that is accepted for a textual change in the test material (Nissi, 2015).

The testers' correct understanding of the problem identified or the change made in the test item is important for accountability issues in testing because in any case of criticism, as the tester of the institution, they will have to explain to the stakeholders why the item is designed in such a way. By means of such interactional actions, the testers also act as a facilitator by coordinating the decision-making activity and directing the trajectory of the talk towards the achievement of some practical end (Barnes, 2007; Nissi, 2015). In doing this, they establish shared understanding of the matter at hand and receive confirmation from the other group members before closing down the sequence since "the emerging text will affect the future of the project members as a collectivity" when the constructed test is administered by the teachers (Nissi, 2015, p.19).

Secondly, in all the three phases in IR interactions, the reviewers display an orientation to their institutional roles as teachers by framing their claims from students' and teaching perspectives. Throughout the IR interactions, while account-giving for a problem identified in the test item in phase PI or evaluating a change in phases SC and RC, the reviewers make reference to past instructional events (Can Daşkın, 2017; Can Daşkın & Hatipoğlu, 2019; You, 2015) and their students' present state of L2 knowledge and skills based on what is taught to them. This is because as teachers, they have an epistemic access to the domain of teaching and are best familiar with their students' L2 knowledge and abilities. This implies the reviewers' epistemic right, responsibility and obligation (Heritage, 2013) to review the item or the changes offered for it with consideration of the teaching and learning context as well as the target group of students' needs and abilities and thus, to create the interplay between teaching and testing. Finally, throughout the talk-in-interaction in all the three segments of extract 10, the reviewer YU orients to his responsibility to read the story carefully before coming to the IR session for the story exam. This is illustrated by his reference to his reading process in extract 10a, his justification for his suggestion with reference to the story book in extract 10b, and his recall of character names from the story as well as his display of knowledge of the suggested character names in extract 10c.

Despite an orientation to the institutional role distribution of the participants, the analysis has also shown that they share the responsibility for problematizing and revising the test item and for their final decision in such task-oriented IR interactions where the product is shared, affecting everybody in the group. Whether it is the reviewer or the tester, the participants all get actively engaged in the IR task, that is in the task of problematizing the item (phase PI), suggesting and deciding on a change to overcome the problem in the item (phase SC) and reviewing the change in the revised item (phase RC). In phase PI, for instance, although it is the reviewers who are responsible for problematizing the item, they invite and follow a process of collaborative inquiry whereby the participants all actively and critically think about the item and make a contribution. Likewise, in phase SC, the participants all have a contribution to make to the change process in the problematized test item. This is evident in their (1) role as suggestion-makers throughout the phase of SC, their constant search and spontaneous formulation of an alternative until they reach a mutually acceptable final resolution, (2) collaborative inquiry into and evaluation of a suggestion from different perspectives, (3) reformulation of a suggested expression in collaborative turn sequences and thus, shared text production (Nissi, 2015), and (4) their joint decision-making

process with regard to the change to be made in the item by, for example, leaving their suggestions open to negotiation and seeking each others' confirmation. As for phase RC, the revised item is re-evaluated and finalized collaboratively by the IR group members as they solicit each others' confirmation for the change and recheck the item with respect to the fulfilment of different testing principles, giving account for it, challenging it where necessary and making further revisions in the item.

In addition, in all the phases, the participants establish mutual agreement and shared understanding collaboratively by using various interactional resources such as collaborative turn completions and sequences including anticipatory completion (Lerner, 1996a, 1996b, 2004; Sert, 2019; Szczepek, 2000; Van Kruiningen, 2013) and choral co-production (Lerner, 2002), additional account, modified repeat (Kim, 2002; Stivers, 2005), candidate understanding (Antaki, 2012; Heritage, 1984b), (re)formulation (Baraldi, 2013; Barnes, 2007; Drew, 2003; Franco & Nielsen, 2018; Heritage & Watson, 1979), and collaborative writing activity. In this way, they argue with one collective voice, not only articulating but also constructing the argument jointly by, for instance, taking up another's line of argument or drawing conclusions from what another says to achieve a shared purpose (Smithson & Diaz, 1996).

Finally, the participants' use of an inclusive language (e.g., inclusive "we") to refer to themselves as members of an organization (Sacks, 1992) is another indication of their shared responsibilities in the IR interaction. It demonstrates that all the co-participants have the epistemic right, responsibility and obligation (Heritage, 2013) to contribute to the IR process and to collaboratively decide on the problem as well as the changes to be made in the item. The participants' use of an institutional self-referring "we" also conveys their stance towards the interactional context in the IR practice (Heritage, 2005). They "invoke an institutional over a personal identity, thereby indicating that they are speaking as representatives, or on behalf, of an organization" and are not personally responsible for the outcomes of the task they are undertaking (Drew & Heritage, 1992, p.30; Sacks, 1992).

The participants' equal active engagement is further evident in their gaze behaviour. It is seen that the participants continuously look at their exam papers only sometimes directing their gazes at each other for a short time. Although a disengaged gaze would make it sanctionable in an ordinary conversation, in such a task-oriented interaction, it does not



indicate disengagement; on the contrary, it signals that the participants are all engaged in the task and actively think about the matter at hand (Goodwin, 1981; Rossano, 2013). As Goodwin (1981) acknowledges, different from ordinary conversation, engagement in a competing activity and looking at an object (i.e., the test material) used for the accomplishment of such an activity (i.e., IR practice) is not a sign of diminished engagement in the conversation. As long as gaze withdrawal is occupied with talk-relevant activities, “the act of disattending each other never emerges for either party as a noticeable, recognizable activity in its own right” and is, in fact, used by the participants as a way of monitoring their actions (Goodwin, 1981, p.107).

In addition, the participants’ non-verbal actions further demonstrate their active engagement as well as their shared responsibility for the IR task and their ownership of the test as the shared product, the outcome of which is important for all of them. They, for instance, scan through the text in search of evidence, search through the textbook or the story book, shift their gazes between the text and the item, turn over the page to look at the reading text or the item, do silent reading, check their notes on the note-taking sheet, underline or move their pencil along the relevant parts of the item/text, circle or point at the words with a pencil, mark the change on their exam papers, and produce a thinking face (Goodwin & Goodwin, 1986). Even when a participant seems to be relatively more passive in the talk-in-interaction, his/her orientation to the test material and sustained gaze at the exam paper show his/her engagement in the task. In this respect, the test material as a shared accessible resource used not as a by-product of activities but as the main object of the meeting, coordinates the participants’ actions throughout the talk-in-interaction in collaboratively accomplishing the task of evaluating, revising and co-constructing the written text in the test material (Nissi, 2015). The visible structure of the participation framework in task-oriented language test IR interactions, hence, represents the collaborative coordination of multiple resources such as talk, recurrent interactional structures, mutually available written materials, gaze, writing and reading in working through a shared task and their role in establishing mutual orientation, achieving a focus of attention on the test items under evaluation and thus, in constituting the ‘jointness’ of the IR activity (Balaman & Sert, 2017; Ford, 1999; Goodwin, 2000, 2007; Lindström & Mondada, 2009). The deployment of such multiple resources indicates that IR practice is a multifaceted, situated and socially constructed activity involving jointly implemented tasks in its each phase (Goodwin 2000, 2007).

The involvement of all the participants in the IR process is especially due to their equal epistemic access to the matters at hand and their constitution of an epistemic community (Heritage, 2013). They assume equal epistemic state and authority since they all put themselves in K+ position as reviewers, teachers and testers who have equal epistemic access to the shared resources and who share a common ground in terms of the testing principles for language test construction (Heritage, 2013). For instance, in extract 4a, since the problem is related to the options deriving from a reading text, which is an accessible source of information for all the participants, they collaboratively investigate the matter at hand by orienting to the reading text in an attempt to prove the problem. They do not just leave it to the problematizer of the item to provide an account but they themselves seek evidence, give an account from different perspectives and present their views with regard to the acceptability of the problem. As a result, the participants do not assert epistemic priority as the first evaluator or suggestion-maker (Heritage & Raymond, 2005) but invite others' contributions to the IR practice, sharing the responsibility for the decision made about the item. In a study on the group dynamics of testcrafting, Davidson and Lynch (2002) also found that the participants did not assume certain roles and their contributions to the discussion changed depending on their teaching, testing and language learning experience without any one of them dominating the group's interaction.

#### **5.4. Implications for Language Test Construction**

The findings of this study have implications for language test construction and the stage of IR in particular. By providing an interactional model for the phases of syllabus-based language test IR practice along with a description of the testing principles as well as the institutional role distribution of the participants oriented to in and through talk, this study is expected to contribute to a number of gaps that have been identified in the literature regarding language test construction. The problems and gaps can be summarized as follows; (1) a number of models including IR as one of the critical stages has been proposed for language test construction (e.g., Bachman & Palmer, 1996; Carr, 2011; Downing, 2006; Fulcher, 2010; Fulcher & Davidson, 2007; Hughes, 1989; Izard, 2005; Milanovic, 2002; Mislevy et al., 2002) and interaction with stakeholders such as teachers has been highlighted especially at this stage (e.g., Bachman & Palmer, 1996; Carr, 2011; Fulcher, 2010; Fulcher & Davidson, 2007; Hughes, 1989; Izard, 2005; Mathew, 2004; Shohamy, 2007; Spaan,

2007); however, the micro-details of interactions that take place with teachers in the process of reviewing language test items have not been studied and presented, (2) the determined procedures and texts on language testing have mainly addressed large-scale standardized testing and focused mostly on the theoretical and statistical aspects of test validity and reliability (Brindley, 2001; Fulcher & Davidson, 2007; Fulcher, 2010; Tan & Turner, 2015), (3) IR practice has generally been described in terms of sub-processes (Fulcher & Davidson, 2007; Fulcher, 2010) and checklists have been suggested to guide those processes (Bachman & Palmer, 1996; Brown, 1996; Brown, 2001; Izard, 2005; Weir, 2005) but there has been no focus on their interactional nature, on whether and how actually reviewers go through those processes in reality and on how they shape the design of test items, (4) there is a number of studies that deal with teacher feedback and the role of interaction with stakeholders in language test construction (e.g., Davidson & Lynch, 2002; Lynch & Davidson, 1994; Green & Andrade, 2010; Green & Hawkey, 2011; Kim et al., 2010; Mathew, 2004; So, 2014; Tan & Turner, 2015) but they mainly focus on large-scale tests and with the exception of So's (2014) and Green and Hawkey's (2011) studies, they do not specifically investigate the stage of IR but rather the design of language tests as a whole (Green & Andrade, 2010; Mathew, 2004; Tan & Turner, 2015) and such language test development stages as determining test specifications (Davidson & Lynch, 2002; Lynch & Davidson, 1994) and item writing (Green & Hawkey, 2011; Kim et al., 2010), and (5) no conversation analytic studies have been carried out on teacher-to-teacher interactions in relation to language test construction in educational settings, and CA studies in language testing have recently been used widely to examine the kinds of interactions co-constructed in different L2 oral proficiency test formats (e.g., Galaczi, 2014; Gan et al., 2008; Kasper & Ross, 2007; Kormos, 1999; Seedhouse, 2013; Seedhouse et al., 2014; Young & He, 1998) as well as the interactional practices that do informal formative assessment in the context of classroom interaction (Can Daşkın & Hatipoğlu, 2019; Heritage & Heritage, 2013).

Against this background, this study is expected to contribute to language test construction and to the development of a better understanding of the IR stage in particular by providing a data-led description of this practice via CA and suggesting interactional resources and actions that are necessary for reviewing and revising tests collaboratively. More specifically, this study presents a vivid picture of IR practice for syllabus-based language tests by showing the micro-analytic details of how test items are problematized, suggestions are provided to change and improve them, changes are reviewed, and how mutual agreement

and understanding as well as joint decision-making is accomplished in and through interaction in the IR meetings with teachers. Such a detailed description demonstrates the close relation between test construction and interaction, which has been emphasized in the testing literature. Since the goal of IR practice is to evaluate the quality and effectiveness of test items in terms of their relevance and usefulness for their intended purpose and make suggestions to improve them (Fulcher, 2010), the findings of this study show how this goal is achieved in and through interaction with EFL teachers.

It has become clear that IR practice cannot simply be realized by following a prescribed step-by-step standard checklists and sub-processes as proposed in the literature since it is found to be a much more delicate work than expected, requiring various professional and interactional skills. Similar to the language test construction process that has been described as iterative (Bachman & Palmer, 1996), IR activity itself also involves an iterative and a dynamic process in practice as evident in the reflexive relationship between the phases of IR interaction and different interactional trajectories and sequential consequences that unfold for the test items under evaluation. Additionally, this study turns the attention to the co-construction of syllabus-based language tests in local contexts, which have been neglected as compared to large-scale testing and thus, brings into view the daily assessment practices of teachers. It contributes to the process-oriented qualitative studies in current language testing studies and more specifically, to conversation analytic language testing studies, which have been limited to language proficiency interactions in speaking tests. In this respect, this study meets the need to go beyond highly abstract theoretical concepts to a process-oriented study of language test construction as it is realized in its own local context as part of teachers' day-to-day work (Davidson & Lynch, 2002; Kim et al., 2010; So, 2014; Shin, 2012).

The findings of this study further illustrate the significance of IR practice in language test construction. They support the views that test items can have various problems gone unnoticed by the item writer even if s/he is experienced and skilled, and IR practice in interaction with colleagues is an effective way of identifying and dealing with those flawed items collaboratively through discussion and negotiation before the test is administered (e.g., Alderson et al., 1995; Bachman & Palmer, 1996; Baranowski, 2006; Brown, 1996; Downing, 2006; Fulcher 2010; Fulcher & Davidson, 2007; Hughes, 1989; Linn, 2006). As expected of item reviewers, the participants in this study ask critical questions and orient to various testing principles, carefully evaluating and discussing the test items together from a

range of dimensions and jointly making decisions about them (Alderson et al., 1995; Hughes, 1989).

This study also shows the importance of teacher involvement and interaction in the IR practice (Alderson et al., 1995; Brindley, 2001; Dikilitaş et al., 2018; Fulcher & Davidson, 2007; Lynch & Davidson, 1994; Xerri & Vella Briffa, 2018). It is seen that teachers can help identify any mismatches between teaching and testing (Lynch & Davidson, 1994) by reflecting on their past instructional events and their students' L2 knowledge and abilities and thus, devise an “ecologically sensitive” test that fulfils their local needs (Fulcher, 2010). Since they are the best sources of information about their students' abilities, knowledge and needs and about the material they have taught, they can easily help devise test tasks that reflect what students have been taught and how the material has been presented and practised in class (Fulcher & Davidson, 2007). This is especially important in the design of syllabus-based classroom language tests for better test validity, an increase in test quality, and positive washback effect on instructional practices (Bachman, 1990; Dikilitaş et al., 2018; Xerri & Vella Briffa, 2018). In this study; however, the teachers are found to review the test items not merely in terms of their instructional practices and their students' L2 abilities but also in relation to a wide range of testing-related issues. They, thus, draw on not only their teaching experience but also their testing and language expertise, making important contributions to the item validation process.

The micro-analytic investigation of language test IR interactions in this study represents a qualitative item validation procedure in locally developing language tests and hence, demonstrates how validity evidence can be obtained from stakeholders through qualitative item reviewing in and through interaction (Downing, 2006; Weir, 2005; Wolf et al., 2008). As explained in Section 5.2, the participants' orientation to the fulfilment or violation of a wide range of testing principles concerning, for example, the content of the item with a consideration of the context of learning and in-class activities, the key of the item, the students' L2 skills and knowledge, the language used, the difficulty level of the item, and testing-related issues such as the plausibility of options and the independency of the item provides different kinds of evidence for validity argument, helping to identify any threats to test validity (Carr, 2011; Fulcher & Davidson, 2007). It is seen that the participants critically evaluate, explain, challenge, and question the claims made about the test items in the IR session by giving account and presenting concrete evidence from the test material (e.g., the

reading text) and the teaching and learning context (McNamara, 2004). The contradictory evidence that arises in validation research is then removed by revising the test item (Davies & Elder, 2005). In this way, the claims about the meaning and use of a test are seen to be supported in a way that is “acceptable to the community of language testers and stakeholders in open discussion, through a process of dialogue and disagreement” (Fulcher & Davidson, 2007, p.21). As a result, the findings of this study have demonstrated that IR practice is part of the argument and evidence for item validity as items are reviewed in terms of, for instance, their content, cognitive demand, and adherence to item writing guidelines and then are either kept as they are, revised, or alternatively discarded depending on the outcome of the IR process (Haladyna & Rodriguez, 2013).

This study supports the findings of studies that have investigated the role of interaction and teacher feedback in language test construction and offers further insights into the way IR practice is specifically implemented in teacher-to-teacher workplace interactions in an educational context. It brings evidence from actual IR interactions for the findings that receiving feedback from colleagues in and through interaction (1) provides useful preliminary evidence for validity argument (Green & Andrade, 2010; Lynch & Davidson, 1994; So, 2014; Tan & Turner, 2015), (2) helps to criticize, reject or modify certain features of texts or items and to eventually come to terms with different aspects of tests through cooperation and open dialogue (Green & Hawkey, 2011; Mathew, 2004), and (3) leads to a productive decision-making process as the participants use certain communication strategies to create and maintain group cohesiveness (Davidson & Lynch, 2002).

On the other hand, different from these studies, this study (1) brings IR practice into the forefront as one of the critical stages of language test construction, (2) goes beyond large-scale tests to in-house syllabus-based language tests that are locally prepared, (3) works with naturally-occurring data and uses CA, which offers the possibility to systematically examine the micro-details of IR interactions from an emic perspective, (4) vividly demonstrates how the sequential unfolding of IR interactions with teachers and the use of certain interactional actions shape the design of test items, affect the decision-making process and contribute to their validation, (5) uncovers the overall structural organization of those interactions and introduces a data-driven model that can be directly used for training purposes, (6) reveals the kinds of testing principles that the participants orient to in and through the IR interactions and how they deal with the violated testing principles, and (7) explains the institutional role

distribution of the participants as it is displayed in the turns-at-talk and discusses tester and reviewer roles and responsibilities that are made relevant in each phase of IR interaction. In this way, besides supporting the existing studies, this study brings insights into their findings and the IR practice in particular.

### **5.5. Implications for EFL Teachers' Professional Competence**

The findings of this study have implications for EFL teachers' professional competence as they reveal (1) the interactional patterns and actions that are necessary for the accomplishment of IR practice and (2) participants' orientations to testing principles as well as their roles and responsibilities as testers and reviewers. In the context of institutional settings, as a central component of the wider social abilities, interactional competence enables people to pursue educational and professional goals and gain access to various institutional worlds (Pekarek Doehler & Petitjean, 2017). Due to the interactional demands and challenges presented by institutions, interactional competencies constitute an important resource to deal with those demands since they are "central instruments for participation, task accomplishment, self-presentation as well as membership" (Pekarek Doehler & Petitjean, 2017, p.10). This means that novice teachers need to adapt their existing interactional competencies and elaborate new ones to deal with a variety of new interactional situations when they move into the world of work (Pekarek Doehler & Petitjean, 2017). To make this possible, research needs to reveal what interactional competencies are solicited, how they are deployed in various contexts and the challenges that emerge for people interacting in those contexts. This study aims to do just this by uncovering the interactional competencies along with participants' orientations to their teaching, testing and language expertise in the context of language test IR activity that is undertaken by teachers in an educational setting via CA. As Seedhouse (2011) indicates, CA treats interactional competence as variable and co-constructed by participants in interaction rather than as fixed and static and offers a means of examining how such competence is co-constructed by participants in particular contexts. Thus, the interactional competencies described in this study using CA can change in different test IR contexts and provide a baseline for subsequent studies and for comparison with their findings.

Considering the three phase of language test IR interactions examined in this study and the sequential environment of each phase, the following interactional actions can be listed to

describe L2 teachers' professional competence in the case of language test IR. The teachers, thus, could consider doing the following in each phase of IR interaction with their colleagues and gain awareness about this practice.

Phase PI:

- Launching a preliminary inquiry into the problem to establish the necessary pre-conditions for the statement of the problem in the item by seeking concrete evidence for it through a question that makes reference to past instructional events or to shared sources, by soliciting confirmation for the answer of the item, or by expressing uncertainty about the correct answer
- Giving account for the problem by presenting evidence from available shared sources, past instructional events, and the students' assumed present state of L2 knowledge and skills and by orienting to testing principles for item writing
- Challenging the problem by presenting counter-evidence from available shared sources, by making counter-claims with respect to the violation of the testing principles addressed by the problematizers of the item, and by the tester referring to the item-writing process
- Managing the disagreement with the problem through a reversal of opinion or conceding activity by discussing the claims made about the item collaboratively and investigating the validity of the arguments and by making certain decisions about the item (e.g., revising, keeping or omitting the item)
- Demonstrating mutual agreement and understanding, and constructing the arguments about the item jointly by using various interactional resources such as full or a partial modified repeat, (re)formulation, additional account or evidence, candidate understanding, candidate evaluation and explanation, collaborative turn completions and sequences including anticipatory completion and choral co-production, as well as reflection on the item writing process
- Elaborating on the problem even when an agreement is displayed with it for a better understanding of the problem
- Performing mitigating actions when announcing the problem in the item by using some interactional resources such as a preface, direct or an embedded question, hedges, mitigating account, and positive assessment as well as when disagreeing with the problem by using agreement-prefaced formulations

Phase SC:

- Suggesting a change to solve the announced problem in the language test item



- Producing a shared text by reformulating a suggested alternative through collaborative turn completions and sequences
- Inviting collaborative inquiry into the suggested alternative by downgrading it through hedges, confirmation requests and its presentation as an example
- Giving account for a suggestion by presenting evidence from available shared sources, past instructional events, and the students' present state of L2 knowledge and skills and by orienting to not only the violated testing principles that were revealed in phase PI but also other testing principles that the offered suggestion would simultaneously fulfil
- Carefully evaluating a suggestion and challenging not only the other participants' suggestions but also one's own suggestion (i.e., disagreeing with 'oneself') by presenting evidence from available shared sources and making reference to the violated testing principles announced in phase PI as well as other testing principles that the suggestion would further violate
- Withdrawing or rejecting a suggestion and proceeding with their search for a better alternative when it is challenged with clear evidence (i.e., agreeing to disagree with a suggestion)
- Playing the devil's advocate and stimulating discussion on an offered suggestion before deciding on it
- Demonstrating mutual agreement and understanding, and discussing the offered suggestions and making a decision about the change jointly by using various interactional resources such as anticipatory completion, additional account or evidence, candidate understanding, collaborative writing activity, full or a partial modified repeat, (re)formulation, candidate explanation, choral co-production, and reflection on the item writing process
- Elaborating on the suggested change even when an agreement is displayed with it for a better understanding of the suggestion
- Initiating a new suggestion-making sequence when one change in the item requires another change
- Writing down the decision as a way of demonstrating their commitment to the proposed action and using the resources of collaborative writing and reading aloud to ensure shared understanding in the group with respect to the change made in the item

#### Phase RC:

- Seeking and checking the group's final confirmation for the revised item often using a tag question that mobilizes support for the change
- Playing the devil's advocate through their confirmation requests by inviting the group to re-evaluate the change with respect to not only the testing principle that is claimed to be violated in the first phase of the IR interaction but also other testing principles for item writing
- Challenging the revised item with reference to testing principles for item writing when they identify a new problem that emerges in the item as a result of its revision in phase SC
- Reinstating phase SC to solve the emerging problem in the revised item
- Confirming the revised item based on evidence often following a collaborative process of account-giving from different perspectives by making reference to the addressed testing principles and their subject knowledge and by presenting concrete evidence from the shared sources
- Elaborating on the revised item even when an agreement is displayed with it to establish shared understanding
- Demonstrating mutual agreement and understanding by using various interactional resources such as additional account or evidence, repetitional response, modified repeat, (re)formulation, candidate explanation, positive assessment, and anticipatory completion

Those interactional actions show that although some competencies are specific to each phase of language test IR interaction (e.g., launching a preliminary inquiry into the problem, suggesting a change, seeking and checking confirmation for the revised item), others are repeated throughout the IR practice in a way that serves the specific goals of each phase. To illustrate, the act of account-giving constitutes an important interactional competence in IR practice since accounts are provided for the problem identified in the item, for the offered suggestion as well as the revised item. They are also used not only to support the problem, suggestion or the change made in the item but also to challenge a claim and display agreement or understanding. Thus, account-giving is a ubiquitous feature of IR interactions (Firth, 1995). Similarly, the act of challenging is a recurring competence that is made relevant when critically discussing the announced problem and evaluating the offered suggestions as well as the revised item before finalizing the item. In this way, multiple

perspectives are considered while completing the task. Additionally, due to the delicacy of the work that is carried out and the significance of the product that is produced in IR interaction, the participants elaborate on the problem, the offered suggestion or the revised item even when an agreement is displayed in the group. They construct the arguments jointly from different perspectives and give a careful thought to the item, the suggested changes, or the revised item by engaging in collaborative inquiry and playing the devil's advocate before making a decision and finalizing the item. They also use similar interactional resources in each phase to establish and demonstrate mutual agreement and understanding (e.g., (re)formulation, anticipatory completion, modified repeat). Furthermore, in phases PI and SC, in managing the disagreement that arises, the participants display an orientation to consensus through a reversal of opinion or abandonment of position by, for example, agreeing to consider the problem after challenging it and then being counter-challenged by the other participants, or agreeing to disagree with an offered suggestion when it is challenged by the other participants with clear evidence (Sharma, 2012; Wasson, 2000).

What is more, it is seen that as part of the participants' interactional competence, they orient to not only transactional goals but also relational goals in the IR practice with colleagues by performing mitigating actions when announcing the problem in the item, disagreeing with the announced problem, or when offering a suggestion while everyone is seeking one to revise the problematized item. The findings also show that throughout the IR interactions, the participants establish mutual orientation and achieve a shared focus of attention on the test items under evaluation by coordinating multiple resources such as talk, recurrent interactional structures, mutually available written materials, gaze, writing and reading. This could be considered as another interactional competence necessary for IR practice especially in terms of constituting 'jointness' in such a task-oriented collaborative workplace interaction. Finally, throughout the language test IR interactions, the participants bring their professional competence to the talk-in-interaction by drawing on their teaching experience as well as their testing and language expertise, which facilitates their use of interactional actions such as account-giving, challenging a claim and displaying agreement or understanding. In doing this, they orient to a wide range of testing principles concerning, for example, the importance and relevance of the content of the item considering their teaching and learning context, the key of the item, students' L2 skills and knowledge, the language used in the item, and testing-related issues such as the plausibility of options.

With regard to the description of language test IR competence revealed in this study, it involves competencies that are required not only for the accomplishment of IR practice but also for task-oriented workplace interactions in general. This is because language test IR practice as a collaborative task involves “not only making conceptual connections [i.e., orientation to testing principles] but also, and crucially, making interactional connections, knowing how to collaboratively construct a task” (Ford, 1999, p.370). The three phases identified in this study with the phase-specific actions of problematizing the item, suggesting a change to revise the problematized item and reviewing the change along with an orientation to testing principles as well as tester and reviewer roles and responsibilities at various stages can be considered within the scope of language test IR competence. On the other hand, the findings with regard to certain interactional actions such as account-giving, challenging, performing mitigating actions, seeking group’s confirmation, managing disagreement, joint suggestion and decision-making as well as those that are used to display mutual agreement and understanding (e.g., candidate understanding, anticipatory completion, (re)formulation) can be considered as components of interactional competence necessary for institutional task-oriented and problem-based interactions in general.

As a result, this study contributes to the understanding of task and goal-oriented workplace interactions in educational settings (e.g., Barnes, 2007; Baker, 1997; Hall & Danby, 2003; Sandlund & Sundqvist, 2019; Van Kruiningen, 2013; Verkuyten, 2000), which are under-researched compared to, for example, business organizations (e.g., Baraldi, 2013; Franco & Nielsen, 2018; Nissi, 2016; Wasson, 2000). It demonstrates not only how language tests are reviewed and revised based on the participants’ teaching, testing and language expertise in particular but also how intersubjectivity and shared understanding as well as collaborative problem-solving and decision-making are accomplished in and through interaction in general. That is, this study unveils the interactional actions that might contribute to joint task completion in professional meetings even though the kind of task may vary in each context depending on the institutional goal. Therefore, despite a focus on a specific institutional task of language test IR and a context-specific research setting with no purpose of making generalizable claims, this study can be said to have implications for any kind of professional contexts where certain tasks need to be accomplished in groups in and through collegial interactions.

## **5.6. Implications for L2 Teacher Education**

The findings of this study have implications for L2 teacher education by presenting a model emerging from the analysis of language test IR interactions where test items are negatively evaluated (See Figure 7). This model can be used to help teachers and student-teachers deal with the demands of such a task in their institutions. Second language assessment constitutes an important part of teachers' expertise (O'Loughlin, 2006). When language teachers start their profession, they are not only expected to teach but also get involved in the test construction process either by writing tests or reviewing exam papers prepared by their colleagues. This means that teachers have multiple roles and need to be prepared for testing practices before they graduate. Even if they are not directly responsible for constructing the tests themselves, they need to understand what constitutes good testing practice so that they can evaluate the claims made for tests by different examining authorities, and from among a range of available tests, select the one that best suits their purpose or moderate tests prepared within their institution (Alderson et al., 1995; Stiggins, 1999). This is also made clear in the description of teachers' language assessment literacy (LAL), which includes the knowledge and ability to design, develop, and evaluate standardized or classroom based tests (Fulcher, 2012). Language test IR competence that is proposed in this study, thus, constitutes an important part of EFL teachers' assessment literacy as well as their institutional responsibilities. They need to develop their ability to evaluate and revise language tests collaboratively in and through interaction with their colleagues to be able to fulfil a part of their testing responsibilities successfully in their professional contexts.

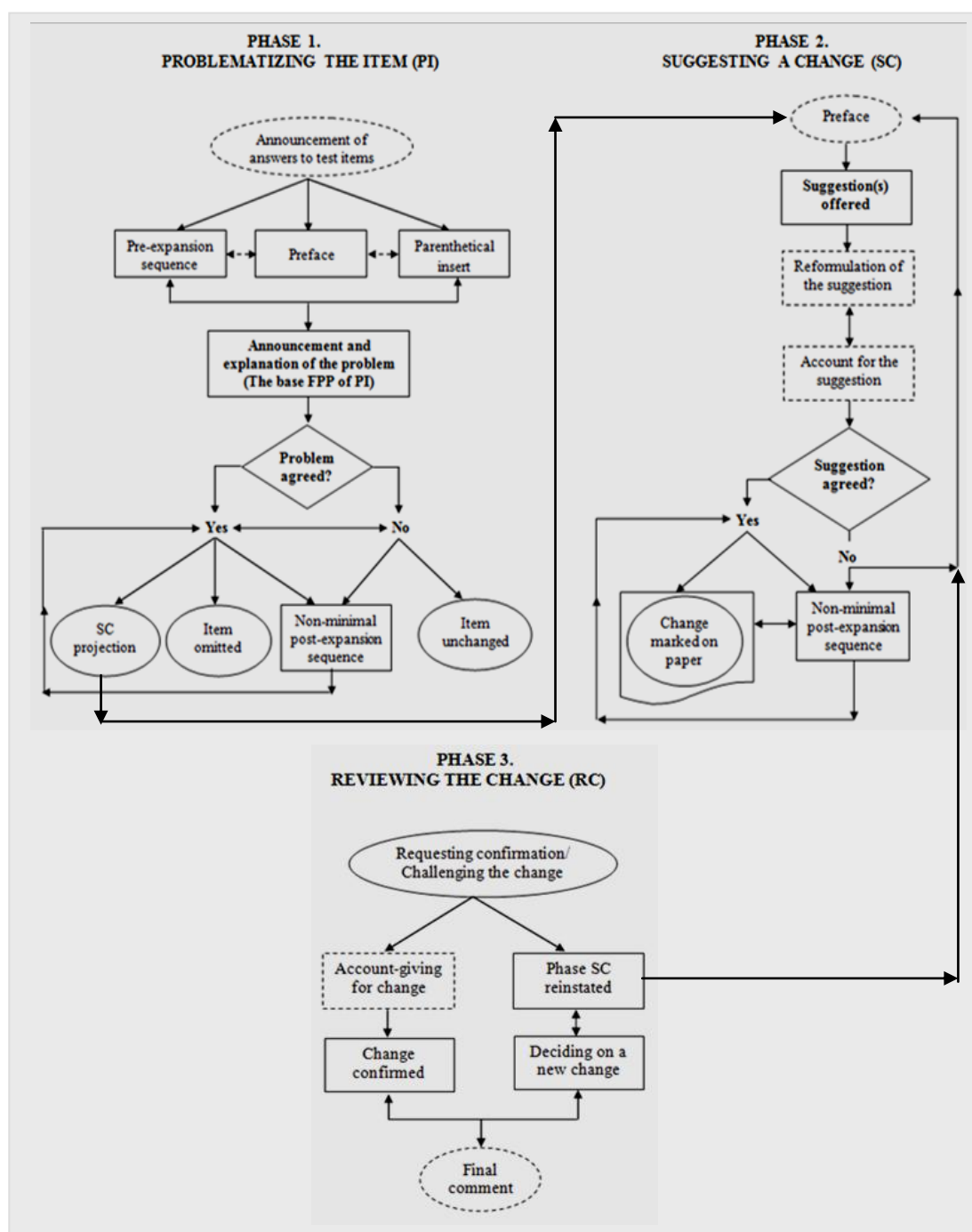


**Figure 7. The Model Emerging from Language Test IR Interactions**

While it is clear that IR activity is an important part of teachers' day-to-day testing practices in language education contexts, there are studies showing that in pre-service teacher education programs, there is a need to devote more time to (a) constructing tests collaboratively through interactions between student-teachers, and (b) critically evaluating,

reviewing and discussing items from various published exams as well as from actual classroom-based tests prepared by student-teachers (e.g., Hatipoğlu, 2010; Kleinsasser, 2005; O'Loughlin, 2006). In a report on European survey of language testing and assessment needs of teachers, teacher trainers and experts, Hasselgreen et al. (2004) revealed that reviewing and writing tests was one of the areas which teachers and experts perceived themselves particularly in need of training. Additionally, studies have identified a need to present student-teachers with the real challenges of testing practices in specific language education settings (e.g., Hatipoğlu, 2010; Stiggins, 2001). In a study on the testing needs of teachers, Fulcher (2012), for instance, found that teachers need a practical how-to guidance that is not necessarily prescriptive in nature and that includes illustrations and practical examples drawn from a range of sources and countries for both classroom and large-scale testing. This means that pre-service teachers need to be familiarized with the actual testing processes that take place at educational institutions both in and outside classrooms and to be shown what they will be doing in relation to testing once they are on the job. This study is expected to meet those needs by presenting the micro-analytic details of language test IR interactions with teachers in a local institutional context, which future teachers will be likely to engage in. Considering the context of this study which is an English preparatory school in a higher education setting in Turkey, language test IR is a routine task commonly performed by teachers at those schools.

Against this background, as Figure 7 illustrates, the interactional language test IR framework that is introduced in this study can be used as a guide for training purposes in monitoring and evaluating language test item reviewer performance in both pre-service (i.e., English Language Testing and Evaluation Courses) and in-service teacher education programs. The findings of this study as summarized in a model in Figure 7 along with the actual samples of talk might be utilized in the form of a testing tutorial as an awareness-raising practice and shared with student-teachers or novice teachers. Those teachers can observe the IR sessions situated in authentic contexts, discuss the micro-analytic investigations of those interactions between teachers, and apply them with their mentor in their own work. In this way, they can learn the skills of IR and gain the interactional competence necessary for accomplishing this practice in line with the needs of their specific contexts. Their awareness can also be raised with respect to the relation between interaction and the practice of syllabus-based language test IR as well as the relevance of the testing knowledge that they gain in their testing courses.

A number of reflective teacher education models has already been proposed for developing student-teachers' classroom interactional competence (e.g., Bozbıyık, 2017; Seedhouse, 2008; Sert, 2015; Walsh, 2003). Those models can be adapted for language test IR practice as a way of helping student-teachers develop their professional and interactional competence for language test IR in particular and for collaborative workplace interactions that they will be likely to engage in with their colleagues in general. To illustrate, Sert (2015) proposed a CA-informed framework named IMDAT, which is a reflective teacher education model that can be implemented in both pre-service and in-service teacher education programs. Such a model integrates conversation analytic findings into teacher education whereby those findings are used as resources and audio-visual materials for training teachers in their pedagogical practices. It requires teachers to not only record their teaching practices but also critically reflect on those recorded interactions in collaboration with their trainers using the principles of CA (Sert, 2015). In this way, student-teachers can be given the opportunity to investigate how teachers implement IR practice in other settings and how they themselves review and revise test items in and through interaction with their peers by using the framework (See Figure 7) proposed for IR interactions in this study as a guideline.

This study suggests an increase in the number of English Language Testing and Evaluation courses in pre-service teacher education programs considering the high amount of practice and time that would be required with the use of CA-informed training models. As Hatipoğlu (2010) showed, a majority of Foreign Language Education departments in state universities in Turkey has only one testing course, which means that university lecturers in those departments are expected to present to the trainee teachers both the theoretical and practical aspects of language testing and assessment in a single course. With the increase in the number of testing courses in EFL departments either as an elective or a must course, a space can be provided in the curriculum of those courses for the use of reflective teacher education models that integrate CA findings on language test IR practice, engage student-teachers in similar IR interactions with their peers, and that stimulate reflection. In this way, they can better translate what they have learnt into real testing environments where they eventually find themselves and testing courses can be made more relevant to the local needs of student-teachers.



## CHAPTER 6

### CONCLUSION

#### 6.0. Introduction

This chapter presents a summary of the findings, conclusions and implications drawn from these findings. Then the limitations of the study and suggestions for further research are discussed.

#### 6.1. Conclusions and Implications

Drawing on a corpus of video-recordings of language test IR interactions (25 hours) between EFL teachers in an English preparatory school at a state university in Turkey, this study has revealed how teachers review language test items prepared for their students in and through interaction with the item writer who is one of the teachers assigned in the testing office of the school. Following the principles of CA, the analysis showed that the interactions mainly focused on the problematic items and they were made up of three recurring phases; (1) problematizing the item (PI), (2) suggesting a change (SC) to revise the problematized item, and (3) reviewing the change (RC) made in the item. These emerging phases make up the overall structural organization of the IR interactions where an item is evaluated negatively by the reviewers. As a result, in those instances, the test items that are made available for evaluation are problematized, revised and reviewed again after being modified. In addition, this study has shown that the discovered phases in IR interactions are locally managed and thus, are subject to variations (Drew & Heritage, 1992). Depending on a range of contingencies, some stages can be left out, the previously completed phases can be reinstated, or the patterns in the phases can be repeated until the goal is reached.

The micro-analysis has also generated models for the overall structural organization of the three emerging phases and revealed recurring interactional actions that lead to mutual understanding and joint decision-making about the test item. Due to the goal orientations of each phase in language test IR interactions, there is a central action that characterizes them:

account-giving for the problem in phase PI, suggestion-making in phase SC, and seeking the group's confirmation for the revised item in phase RC. While some of the uncovered interactional actions are specific to each phase (e.g., launching a preliminary inquiry into the problem in phase PI; suggesting a change, producing a shared text, disagreeing with 'oneself' and writing down the decision in phase SC; seeking confirmation for the revised item in phase RC), others are repeated throughout the IR practice in a way that serves the specific goals of each phase.

The acts of account-giving, challenging, and demonstrating mutual agreement and understanding via the use of various interactional resources (e.g., anticipatory completion, modified repeat, (re)formulation) are seen in all the phases with important sequential consequences and implications for the test item under evaluation. In employing those actions, the participants bring their professional competence to the talk-in-interaction by drawing on their teaching experience (e.g., past instructional events, students' L2 knowledge and abilities) as well as their testing and language expertise and by presenting evidence from available shared sources. In this respect, this study has shown that in IR interactions, accounts based on such grounds can be used for multiple purposes such as to (1) support the problem, suggestion or the change made in the item, (2) challenge a claim, (3) display agreement or to ensure shared understanding in the group. This illustrates that accounts have a proactive, creative and problem-solving function in such collaborative workplace interactions as language test IR interactions (Firth, 1995).

These actions of account-giving, challenging, and demonstrating mutual agreement and understanding lead to the collaborative discussion and negotiation of the problem, the suggested changes, or the revised item from multiple perspectives, which results in a better understanding of the raised issues in a problem-oriented context. The participants construct their arguments jointly (Smithson & Diaz, 1996) and give a careful thought to the item, the suggested changes, or the revised item before ending the IR interaction for the item in question. In relation to this, the participants invite collaborative inquiry into the topics of discussion (i.e., the identified problem, suggested alternatives, changes made in the item) by downgrading the problem or the suggestion in phases PI and SC (Koester, 2002), playing the devil's advocate in phases SC and RC (Davidson & Lynch, 2002; Leyland et al., 2016; MacDougall & Baum, 1997; Peercy et al., 2015), launching a preliminary inquiry in a pre-expansion sequence in phase PI (Pomerantz & Heritage, 2013), and challenging the matter at

hand as encountered in all the phases. Additionally, the participants tend to elaborate on the problem, the offered suggestion or the revised item in a non-minimal post-expansion sequence (Schegloff, 2007; Stivers, 2013) even when an agreement is displayed in the group.

Regarding the management of disagreement that arises in phases PI and SC as a result of challenging one another's claims, it is seen that the participants display an orientation to consensus through a reversal of opinion or abandonment of position by, for example, agreeing to consider the problem after challenging it and then being counter-challenged by the other participants, or agreeing to disagree with an offered suggestion when it is challenged by the other participants with clear evidence (Sharma, 2012; Wasson, 2000). The participants deploy certain pre-conceding activities (e.g., redesigning a prior account, giving further account and evidence as a counter-challenge) to exit from disagreements so that some subsequent action can be performed in relation to the task at hand (Sharma, 2012). As a result of the discussion and evaluation of the matters at hand through account-giving and challenging, certain decisions are made about the test items jointly in each phase. In phase PI, the IR group decides to revise, omit or keep the item as it is, or they end up redesigning the announced problem. In phase SC, the participants reject or accept an offered suggestion or they themselves withdraw their own suggestion right after proposing it. In phase RC, the group either confirms the change or agrees with the challenge presented towards the change, reinstating phase SC to solve the emerging problem in the revised item. This shows that language test IR interactions require continuous reviewing and editing with a careful and critical eye since a major change might lead to other changes to make the revision of the item complete.

Different preferences are displayed for disagreement in phases PI and SC. While it is often treated as a dispreferred response in phase PI, it is found to be acceptable and unmarked in phase SC with no extra interactional work. This can be explained by their goal orientations and the evaluative practice being an integral part of task accomplishment in IR interactions (Koester, 2006; Lindström & Mondada, 2009). In phase PI where the aim is to carefully evaluate the item and identify any potential problems in the item, challenges are presented towards the negative evaluation of the test item being reviewed. On the other hand, in phase SC where the aim is to seek the best possible solution in revising the problematized test item, challenges constitute the negative evaluation itself with respect to an offered

suggestion. The way disagreement is managed in those phases demonstrates that negative evaluations which are formulated through clear accounts and evidence either in the base FPP (phase PI) or SPP (in response to a suggestion in phase SC) are expected to be met with the group's agreement for the accomplishment of the task at hand. As a result, when the participants' negative evaluation of the test item is challenged in phase PI where it is usually performed with a form of mitigation (i.e., agreement-prefaced disagreement), the group members often pursue agreement with the problem identified in the test item through some interactional work (e.g., sequence expansions), orienting to disagreement with the announced problem (i.e., negative evaluation of the item) as a dispreferred response. On the contrary, when an offered suggestion is challenged in phase SC, the group members often do not contest the negative evaluation with regard to their suggestion and continue their search for the best possible solution, showing no discomfort with disagreeing responses to their suggestions.

On the other hand, the findings have shown that despite different preferences for disagreement in phases PI and SC, the act of challenging and display of an oppositional stance can change the course of action in different ways particularly when the participants agree with the disagreements and accept the challenges presented towards a claim or a suggestion. In other words, disagreements can affect the decisions made about the test items considerably and play an important role in establishing mutual understanding in the group. This means that discussing a matter at hand with multiple perspectives is important in language test IR interactions as long as it is managed against a background of consensus and shared understanding (Myers, 1998). Disagreement is, thus, found to be a task-supporting activity that avoids poorly judged decisions in institutional settings (Angouri, 2012; Kangasharju, 2002).

The findings of this study have further revealed that the participants orient to not only transactional goals but also relational goals (Koester, 2006) in the IR practice with colleagues by performing mitigating actions (e.g., hedges, mitigating account, delaying devices, agreement-prefaced disagreements) when announcing the problem in the item, disagreeing with the announced problem, or when offering a suggestion while everyone is seeking one to revise the problematized item. In this way, they display not only other-attentiveness but also take into consideration issues of face (Brown & Levinson, 1978; Goffman, 1955) and self-other relations (Heritage, 1988).

The participants are found to display an orientation to various testing principles in and through talk as they collaboratively review and revise the test items. They orient to a wide range of testing principles described in the literature for item writing in general and for selected response (i.e., multiple-choice, matching) or limited production (i.e., fill-in-the-blanks) item formats in particular (e.g., Carr, 2011; Fulcher, 2010; Haladyna & Rodriguez, 2013; Harris, 1969; Hatipoğlu, 2009; Heaton, 1990; Kehoe, 1995). The participants make reference to testing principles most notably while giving account for the matter at hand and challenging it. In those instances, the oriented testing principles are claimed to be either violated or fulfilled by the test item and the participants move towards solving the problem if the item is found to violate a testing principle. The testing principles that are addressed concern issues such as the importance and relevance of the item content considering the teaching and learning context, the key of the item (e.g., “only one correct answer” principle), students’ L2 skills and knowledge (e.g., “level of difficulty” principle), the language used in the item, and testing-related issues such as the plausibility of options, the simplicity of vocabulary in a reading text, sufficiency of context, and the independency of the item. This demonstrates that the participants analyze and evaluate the language test items from various dimensions, drawing on not only their teaching experience but also their testing and language expertise. In this way, they invoke larger principles and base their claims on reasonable and professional grounds (Waring, 2017).

Throughout the phases of language test IR interactions, an orientation is displayed to the institutional role distribution of the participants. In phase PI, the roles that are made relevant in and through talk are more distinctive because the product being reviewed is ‘owned’ by the tester and it is the place where the teachers have the primary rights to evaluate the test item as their main goal in the IR interaction (Heritage, 2013). Therefore, in phase PI, the reviewers problematize the test item by giving account and evidence and the testers employ such actions as announcing the expected correct answer, pre-emptively giving an elaborative account for the expected correct answer to provide insight into the projected problem, and challenging the problem as they become accountable for the problem in their product. However, this role distribution disappears slightly in the following phases since the group reproduces the item collaboratively as the product of shared work. In those phases, their roles and responsibilities are not as stable as they are in phase PI, being made relevant in more than one phase. To illustrate, in phases PI and SC, the reviewers reflect on the test-taking process (Alderson et al., 1995) whereas the testers reflect on the item-writing process.

It is also seen that in all the three phases, the reviewers orient to their institutional roles as teachers by framing their claims from students' and teaching perspectives and creating the interplay between teaching and testing while the testers orient to their responsibility to ensure correct and shared understanding of the problem announced in phase PI or of the accepted changes in phases SC and RC.

Despite an orientation to the institutional role distribution of the participants, there is clear evidence for their shared responsibility and equal engagement in the IR task. They constitute 'jointness' in such a task-oriented collaborative workplace interaction, which is most evident in their collaborative interactional actions, use of an inclusive language (e.g., inclusive "we") (Sacks, 1992), gaze behaviour, and non-verbal actions (e.g., scanning through the text, circling or pointing at the words with a pencil, producing thinking face, shifting gaze between the text and the item). They establish mutual orientation and achieve a shared focus of attention on the test items under evaluation by coordinating multiple resources such as talk, recurrent interactional structures, mutually available written materials, gaze, writing and reading (Balaman & Sert, 2017; Ford, 1999; Goodwin, 2000, 2007; Lindström & Mondada, 2009). The involvement of all the participants in the IR process is especially due to their equal epistemic access to the matters at hand and their constitution of an epistemic community (Heritage, 2013).

The findings of this study have implications for language test construction and the stage of IR in particular by contributing to a number of gaps that have been identified in the literature. Firstly, in language test construction models, IR is included as one of the critical stages and at this stage, interaction with stakeholders such as teachers has been highlighted (e.g., Bachman & Palmer, 1996; Carr, 2011; Fulcher, 2010). Although they do not specifically focus on the IR practice, there is also a number of studies that deal with teacher feedback and the role of interaction with stakeholders in language test construction (e.g., Davidson & Lynch, 2002; Green & Hawkey, 2011; Kim et al., 2010; So, 2014; Tan & Turner, 2015). However, to the best knowledge of the researcher, despite the importance given to interaction and IR practice in language test construction, the micro-details of interactions that take place with teachers in the process of constructing and reviewing language test items have not been studied and presented. This study addresses this gap by (1) working with naturally-occurring data and using CA, which offers the possibility to systematically examine the micro-details of IR interactions from an emic perspective, and

(2) providing a model for the phases of syllabus-based language test IR interactions along with a description of the recurring interactional actions as well as the testing principles and the institutional role distribution of the participants oriented to in and through talk. Such a detailed micro-analytic description via CA demonstrates the close relation between test construction and interaction.

Secondly, this study turns the attention to the co-construction of syllabus-based language tests in local contexts, which have been neglected as compared to large-scale testing and thus, brings into view the daily assessment practices of teachers. It contributes to the process-oriented qualitative studies in current language testing studies and more specifically, to conversation analytic language testing studies, which have mainly dealt with language proficiency interactions in speaking tests (e.g., Galaczi, 2014; Seedhouse, 2013; Young & He, 1998). In this respect, this study meets the need to go beyond highly abstract theoretical concepts to a process-oriented study of language test construction as it is realized in its own local context as part of teachers' day-to-day work. Thirdly, this study has shown that IR practice cannot simply be realized by following a prescribed step-by-step standard checklists and sub-processes as proposed in the literature (e.g., Bachman & Palmer, 1996; Brown, 2001; Fulcher & Davidson, 2007) since it is found to involve an iterative and a dynamic process in practice as evident in the reflexive relationship between the phases of IR interaction and different interactional trajectories and consequences that unfold for the test items under evaluation.

The findings of this study illustrate the significance of IR practice in language test construction by supporting already established views in language test construction texts and studies as well as offering new insights into the way IR practice is specifically implemented in teacher-to-teacher workplace interactions. Firstly, they support the views that test items can have various problems gone unnoticed by the item writer even if s/he is experienced and skilled, and IR practice in interaction with colleagues is an effective way of identifying and dealing with those flawed items collaboratively through discussion and negotiation before the test is administered (e.g., Alderson et al., 1995; Bachman & Palmer, 1996; Baranowski, 2006; Brown, 1996; Downing, 2006; Fulcher 2010; Fulcher & Davidson, 2007; Hughes, 1989; Linn, 2006). Secondly, this study brings evidence from actual IR interactions for the finding that receiving feedback from colleagues in and through interaction provides useful preliminary evidence for validity argument (Green & Andrade, 2010; Lynch & Davidson,

1994; So, 2014; Tan & Turner, 2015). It, thus, demonstrates vividly how validity evidence can be obtained from stakeholders through IR practice in and through interaction, representing a qualitative item validation procedure (Downing, 2006; Weir, 2005; Wolf et al., 2008).

Thirdly, this study illustrates the importance of teacher involvement and interaction in the IR practice (Alderson et al., 1995; Brindley, 2001; Dikilitaş et al., 2018; Fulcher & Davidson, 2007; Xerri & Vella Briffa, 2018). It is seen that teachers can help identify any mismatches between teaching and testing (Lynch & Davidson, 1994) by reflecting on their past instructional events and their students' L2 knowledge and abilities and thus, devise an “ecologically sensitive” test that fulfils their local needs (Fulcher, 2010). They also review the test items in relation to a wide range of testing-related issues, making important contributions to the item validation process. As a result, this study provides new insights and leads to the development of a better understanding of the IR stage as a social practice in language test construction by demonstrating how the sequential unfolding of IR interactions with teachers and the use of certain interactional actions shape the design of test items, affect the decision-making process and contribute to their validation.

The findings of this study also have implications for EFL teachers' professional competence by describing the interactional patterns and actions that are necessary for the accomplishment of IR practice along with the participants' orientations to testing principles as well as their roles and responsibilities as testers and reviewers. In the context of institutional settings, as a central component of the wider social abilities, interactional competence enables people to pursue educational and professional goals and gain access to various institutional worlds (Pekarek Doehler & Petitjean, 2017). By uncovering the interactional competencies in the context of language test IR activity that is undertaken by teachers in an educational setting via CA, this study is expected to help novice teachers deal with the demands of new interactional situations when they move into the world of work (Pekarek Doehler & Petitjean, 2017). Besides, despite a focus on a specific institutional task of language test IR and a context-specific research setting, this study can be said to have implications for professional contexts where certain tasks need to be accomplished jointly in groups in and through collegial interactions. Therefore, this study at the same time contributes to the understanding of teachers' task and goal-oriented workplace interactions in educational settings (e.g., Barnes, 2007; Van Kruiningen, 2013; Verkuyten, 2000).



Lastly, the findings of this study have implications for L2 teacher education as it presents a model emerging from the analysis of the language test IR interactions where items were problematized (See Figure 7). This model can be used to help teachers and student-teachers deal with the demands of such a task in their institutions. Language test IR competence that is proposed in this study constitutes an important part of EFL teachers' assessment literacy as well as their institutional responsibilities. They need to develop their ability to evaluate and revise language tests collaboratively in and through interaction with their colleagues to be able to fulfil a part of their testing responsibilities successfully in their professional contexts. While it is clear that IR activity is an important part of teachers' day-to-day testing practices in language education contexts, there are studies showing that in pre-service teacher education programs, there is a need to devote more time to the co-construction and critical evaluation of tests through interactions between student-teachers and a need to present them with the real challenges of testing practices in specific language education settings (e.g., Hatipoğlu, 2010; Kleinsasser, 2005; O'Loughlin, 2006; Stiggins, 2001). This study is expected to meet those needs by presenting the micro-analytic details of language test IR interactions with teachers in a local institutional context, which future teachers will be likely to engage in.

The interactional language test IR framework that is introduced in this study can be used as a guide for training purposes in monitoring and evaluating language test item reviewer performance in both pre-service (i.e., English Language Testing and Evaluation Courses) and in-service teacher education programs. The findings of this study along with the actual samples of talk might be utilized in the form of a testing tutorial as an awareness-raising practice and shared with student-teachers or novice teachers. In doing this, some reflective teacher education models (e.g., Sert, 2015) can be adapted and used as a way of helping student-teachers develop their competence for language test IR in particular and for collaborative workplace interactions in general. In this way, they can not only understand the role of interaction in language testing practices but also see when and how teachers apply their testing expertise in institutional contexts and thus, find out the relevance of what they learn in testing courses. This study suggests an increase in the number of English Language Testing and Evaluation courses in pre-service teacher education programs in Turkey. In this way, a space can be provided in the curriculum of those courses for the use of reflective teacher education models that integrate CA findings on language test IR practice, engage

student-teachers in similar IR interactions with their peers, and that stimulate reflection, making testing courses more relevant to their local needs.

## **6.2. Limitations**

Like other research methods, which have their own limitations, the use of a qualitative research method presents some limitations for this study. Firstly, the data comes from a specific EFL testing context in a particular institution at tertiary level although language test IR sessions carried out by different groups of teachers and testers over an extensive period of time were recorded for the representativeness of the context. The field would, thus, benefit from further research in different educational contexts where teachers review and revise test items in and through interaction with their colleagues. Although this study does not aim to make generalizable claims, it is also hoped that the findings and general conclusions will be relevant and transferable to the syllabus-based language test construction practices in other language teaching programs in Turkey and around the world where tests are locally prepared and classroom teachers review and co-construct language tests collaboratively.

Secondly, this study has some technical limitations. Since the object of the video-recordings was only small groups (often dyads and triads) working together on a task around a table, one camera was set up and two audio-recorders were placed on two ends of the table around which the participants were seated. The camera was also positioned in such a way that it recorded the entire participation framework. However, in only a few recordings, some of the participants' faces could not be captured at certain time intervals as they changed their positions and the researcher was not present during the recording to make the necessary adjustments in the view of the camera. In addition, despite the availability of sufficient number of microphones for such small group interactions, some of the talk could not be captured at times due to all the participants' active engagement and simultaneous actions in the task as well as their gazes and faces being mostly positioned on the test material occupied by talk-relevant activities, which made it difficult to follow their mouth movements.

Thirdly, there could be limitations in the transcriptions of the data since there is never a complete transcript. Attempts were made to include as many details as possible in the transcripts by viewing and listening to segments of the recorded data repeatedly and by

marking the multi-modal actions that especially represented the participants' task orientations. Some of the extracts were also presented at Conversation Analysis data sessions at HUMAN where the researchers gave some suggestions about the transcripts. However, the transcription process is so dynamic that it is open to even further suggestions.

### **6.3. Further Research**

This study has investigated syllabus-based test IR interactions in an EFL context as they are employed by EFL teachers at tertiary level in Turkey. Further research is needed to expand the findings of this study with the analysis of classroom test IR practice in diverse contexts. A comparative approach that can take this study further can be encouraged by examining IR interactions (1) in similar settings (i.e., English preparatory schools in universities) in Turkey and in other countries, which might reveal cross-linguistic and cross-cultural variations in the way the practice is implemented, (2) at different levels of education such as primary and secondary education, and (3) for the syllabus-based tests of other subjects such as maths and science as well as foreign/second languages other than English and L1 courses. Since this study is the first to offer an interactional framework for syllabus-based language test IR practice using CA, it can function as a springboard for all the other studies interested in the interplay between interaction and language test construction. With the accumulation of empirical findings in time with respect to the interactional organization of IR practice in similar or different educational environments, setting-specific features, similar instances as well as deviant cases can be discovered and the underlying pattern of IR practice revealed in this study can be expanded and built on (Arminen, 2005; Drew, 2003; Gardner, 2004; Heritage & Clayman, 2010; Ten Have, 2007).

Since the reviewers in this study were all experienced teachers, follow-up studies can also focus on how novice teachers perform in language test IR interactions with their colleagues. Their IR performance and LAL development in and through interaction can be tracked through longitudinal studies. In addition, the same kind of data involving participants with different years of experience in teaching and testing can be examined from different dimensions such as politeness and power relations. Another line of research would be to investigate different kinds of task-oriented teacher-to-teacher workplace interactions in educational settings. This would also make it possible to compare the findings of this study in terms of the competencies required of teachers in dealing with the interactional demands

of their institutions. Lastly, CA-informed studies or projects that apply the findings of this study in pre-service or in-service teacher education programs can be carried out in applied linguistic research contexts with implications for task design and material development for English Language Testing and Evaluation courses and for professional development. This would contribute to the understanding of the relevance of this study for training purposes.



## REFERENCES

- Adelswärd, V., & Öberg, B. M. (1998). The function of laughter and joking in negotiation activities. *International Journal of Humor Research*, 11(4), 411-429.
- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. New York: Cambridge University Press.
- Allwright, D., & Bailey, K. M. (1991). *Focus on the language classroom*. Cambridge: Cambridge University Press.
- Angouri, J. (2012). Managing disagreement in problem solving meeting talk. *Journal of Pragmatics*, 44(12), 1565-1579.
- Antaki, C. (2011). Six kinds of applied conversation analysis. In C. Antaki (Ed.), *Applied conversation analysis: Intervention and change in institutional talk* (pp. 1-14). Basingstoke, UK: Palgrave Macmillan.
- Antaki, C. (2012). Affiliative and disaffiliative candidate understandings. *Discourse Studies*, 14(5), 531-547.
- Arminen, I. (2005). *Institutional interaction: Studies of talk at work*. Aldershot, UK: Ashgate.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Baker, C. D. (1997). Ticketing rules: Categorization and moral ordering in a school staff meeting. In S. Hester & P. Eglin (Eds.), *Culture in action: Studies in membership categorization analysis* (Vol. 4, pp. 77-98). Washington, DC: University Press of America.

- Balaman, U. (2016). A conversation analytic study on the development of interactional competence in English in an online task-oriented environment. Unpublished doctoral thesis. Hacettepe University, Ankara.
- Balaman, U. (2019). Sequential organization of hinting in online task-oriented L2 interaction. *Text & Talk*, 39(4), 511-534.
- Balaman, U., & Sert, O. (2017). The coordination of online L2 interaction and orientations to task interface for epistemic progression. *Journal of Pragmatics*, 115, 115-129.
- Baraldi, C. (2013). Forms of decision making: Gatekeeping and dialogic coordination in CISV organizational meetings. *Journal of Business Communication*, 50(4), 339-361.
- Baranowski, R. A. (2006). Item editing and editorial review. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 349-358). New Jersey: Lawrence Erlbaum.
- Barnes, R. (2007). Formulations and the facilitation of common agreement in meetings talk. *Text & Talk*, 27(3), 273-296.
- Bozbiyık, M. (2017). *The implementation of VEO in an English language education context: A focus on teacher questioning practices*. Unpublished Master's Thesis, Gazi University, Ankara.
- Brindley, G. (1997). Assessment and the language teacher: Trends and transitions. *The Language Teacher Online*, 21(9), 37-39.
- Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 126-136). Cambridge: Cambridge University Press.
- British Council. (2015). *The state of English in higher education in Turkey: A baseline study*. Ankara: Yorum Basın Yayın.
- Brookhart, S. M. (1999). The art and science of classroom assessment: The missing part of pedagogy. *ASHE-ERIC Higher Education Report*, 27(1).
- Brown, J. D. (1996). *Testing in language programs*. New Jersey: Prentice Hall Regents.

- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy*. New York: Longman.
- Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing*, 20(1), 1-25.
- Brown, P., & Levinson, S. C. (1978). Universals in language usage. In E. N. Goody (Ed.), *Questions and politeness: Strategies in social interaction* (pp. 56-311). Cambridge: Cambridge University Press.
- Buck, G. (2009). Challenges and constraints in language test development. In J. C. Alderson (Ed.), *The politics of language education: Individuals and institutions* (pp. 166-184). Bristol: Multilingual Matters.
- Can Daşkın, N. (2017). A conversation analytic study of reference to a past learning event in L2 classroom interaction: Implications for informal formative assessment. Unpublished doctoral dissertation. Middle East Technical University, Ankara.
- Can Daşkın, N., & Hatipoğlu, Ç. (2019). Reference to a past learning event in teacher turns in an L2 instructional setting. *Journal of Pragmatics*, 142, 16–30.
- Carr, N. T. (2011). *Designing and analyzing language tests: Oxford handbooks for language teachers*. Oxford, UK: Oxford University Press.
- Chapelle, C. A. (2012). Conceptions of validity. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 21-33). London/New York: Routledge.
- Davidson, F. & Lynch, B. K. (2002). *Testcraft: A teacher's guide to writing and using language test specifications*. New Haven/London: Yale University.
- Davies, A., & Elder, C. (2005). Validity and validation in language testing. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 795-814). New Jersey: Lawrence Erlbaum.
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 305-334.

- Dikilitaş, K., Keşli Dollar, Y., & Mede, E. (2018). English language teachers' perceptions of involvement in high-stakes testing in Turkey. In D. Xerri & P. V. Briffa (Eds.), *Teacher involvement in high-stakes language testing* (pp. 233-244). Cham: Springer.
- Douglas, D. (2010). *Understanding language testing*. London/New York: Routledge.
- Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 3-26). New Jersey: Lawrence Erlbaum.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
- Drew, P. (2003). Comparative analysis of talk-in-interaction in different institutional settings: A sketch. In P. Glenn, C. D. LeBaron & J. Mandelbaum (Eds.), *Studies in language and social interaction: In honor of Robert Hopper* (pp. 293-308). Mahwah, NJ: Erlbaum.
- Drew, P. (2005). Conversation analysis. In K. L. Fitch & R. E. Sanders (Eds.), *Handbook of language and social interaction* (pp. 71-102). Mahwah, NJ: Lawrence Erlbaum.
- Drew, P. (2013). Turn design. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 131-149). West Sussex: Blackwell.
- Drew, P., & Heritage, J. (1992). Analyzing talk at work: An introduction. In P. Drew & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 3-65). Cambridge: Cambridge University Press.
- Erickson, G., & Åberg-Bengtsson, L. (2012). A collaborative approach to national test development. In D. Tsagari & I. Csepes (Eds.), *Collaboration in language testing and assessment* (Vol. 26, pp. 93-108). Frankfurt: Peter Lang.
- Firth, A. (1995). 'Accounts' in negotiation discourse: A single-case analysis. *Journal of Pragmatics*, 23(2), 199-226.
- Ford, C. E. (1999). Collaborative construction of task activity: Coordinating multiple resources in a high school physics lab. *Research on Language and Social Interaction*, 32(4), 369-408.



- Fox, B. A., Benjamin, T., & Mazeland, H. (2013). Conversation analysis and repair organization: Overview. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1094-1097). Oxford: Wiley-Blackwell.
- Franco, L. A., & Nielsen, M. F. (2018). Examining group facilitation in situ: The use of formulations in facilitation practice. *Group Decision and Negotiation*, 27, 735-756.
- Freeman, D. (2009). The scope of second language teacher education. In A. Burns & J.C. Richards (Eds.), *The Cambridge guide to second language teacher education* (pp. 11-19). New York: Cambridge University Press.
- Fulcher, G. (2010). *Practical language testing*. London: Hodder Education.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. London/New York: Routledge.
- Galaczi, E. D. (2008). Peer-peer interaction in a speaking test: The case of the First Certificate in English examination. *Language Assessment Quarterly*, 5(2), 89-119.
- Galaczi, E. D. (2014). Interactional competence across proficiency levels: How do learners manage interaction in paired speaking tests?. *Applied Linguistics*, 35(5), 553-574.
- Gan, Z. (2010). Interaction in group oral assessment: A case study of higher-and lower-scoring students. *Language Testing*, 27(4), 585-602.
- Gan, Z., Davison, C., & Hamp-Lyons, L. (2008). Topic negotiation in peer group oral assessment situations: A conversation analytic approach. *Applied Linguistics*, 30(3), 315-334.
- Gardner, R. (2004). Conversation analysis. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 262-284). Oxford: Blackwell Publishing.
- Goetsch, D. L., & Davis, S. (2014). *Quality management for organizational excellence: Introduction to total quality* (7th ed.). Upper Saddle River, NJ: Pearson.

- Goffman, E. (1955). On face-work: An analysis of ritual elements in social interaction. *Psychiatry: Journal for the Study of Interpersonal Processes*, 18(3), 213-231.
- Goffman, E. (1981). Response cries. In E. Goffman (Ed.), *Forms of talk* (pp. 78-122). Oxford: Blackwell.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Goodwin, C. (1987). Forgetfulness as an interactive resource. *Social psychology quarterly*, 50(2), 115-131.
- Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32(10), 1489-1522.
- Goodwin, C. (2003). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 217-41). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Goodwin, C. (2007). Participation, stance and affect in the organization of activities. *Discourse & Society*, 18(1), 53-73.
- Goodwin, M. H., & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica*, 62(1-2), 51-76.
- Goodwin, C., & Heritage, J. (1990). Conversation analysis. *Annual Review of Anthropology*, 19, 283-307.
- Green, B. A., & Andrade, M. S. (2010). Guiding principles for language assessment reform: A model for collaboration. *Journal of English for Academic Purposes*, 9(4), 322-334.
- Green, A., & Hawkey, R. (2011). Re-fitting for a different purpose: A case study of item writer practices in adapting source texts for a test of academic reading. *Language Testing*, 29(1), 109-129.
- Gronlund, N. E. (1985). *Measurement and evaluation in teaching* (5th ed.). New York: Macmillan.

- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. New York, NY: Routledge.
- Hall, G., & Danby, S. (2003). Teachers and academics co-constructing the category of expert through meeting talk. Paper presented to the Australian Association for Research in Education/New Zealand Association for Research in Education Annual Conference, Auckland.
- Hamid, M. O. (2015). Policies of global English tests: Test-takers' perspectives on the IELTS retake policy. *Discourse: Studies in the Cultural Policies of Education*, 37(3), 472-487.
- Hamp-Lyons, L. (2000). Social, professional and individual responsibility in language testing. *System*, 28(4), 579-591.
- Hamp-Lyons, L. (2007). The impact of testing practices on teaching. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 487-504). New York: Springer.
- Harris, D.P. (1969). *Testing English as a second language*. New York: McGraw Hill.
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). European survey of language testing and assessment needs. *Report: part one—general findings*. University of Bergen.
- Hatipoğlu, Ç. (2009). Writing and evaluating multiple choice Items. A training module prepared for the E-INSET Project.
- Hatipoğlu, Ç. (2010). Summative evaluation of an English Language Testing and Evaluation course for future English language teachers in Turkey. *English Language Teacher Education and Development (ELTED)*, 13, 40-51.
- Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Research Journal*, 4(2), 111-128.

- Hayashi, M. (2013). Turn allocation and turn sharing. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 167-190). West Sussex: Blackwell.
- Heath, C., & Luff, P. (2013). Embodied action and organizational activity. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 283-307). West Sussex: Blackwell.
- Heaton, J. B. (1990). *Writing English language tests*. London & New York: Longman.
- Henning, G. (1987). *A guide to language testing: Development, evaluation, and research*. Boston, MA: Heinle & Heinle.
- Hepburn, A., & Bolden, G. B. (2013). The conversation analytic approach to transcription. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 57-76). West Sussex: Blackwell.
- Heritage, J. (1984a). *Garfinkel and ethnomethodology*. New York, NY: Polity.
- Heritage, J. (1984b). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 299-345). Cambridge: Cambridge University Press.
- Heritage, J. (1988). Explanations as accounts: A conversation analytic perspective. In C. Antaki (Ed.), *Analysing everyday explanation: A casebook of methods* (pp. 127-144). London: Sage.
- Heritage, J. (1997). Conversation analysis and institutional talk: Analyzing data. In D. Silverman (Ed.), *Qualitative analysis: Issues of theory and method* (pp. 161-182). London: Sage.
- Heritage, J. (2002). *Oh*-prefaced responses to assessments: A method of modifying agreement/disagreement. In C. E. Ford, B. A. Fox & S. A. Thompson (Eds.), *The language of turn and sequence* (pp. 196-224). Oxford, UK: Oxford University Press.
- Heritage, J. (2005). Conversation analysis and institutional talk. In K. L. Fitch & R. E. Sanders (Eds.), *Handbook of language and social interaction* (pp. 103-148). Mahwah, NJ: Lawrence Erlbaum.

- Heritage, J. (2011). Territories of knowledge, territories of experience: Empathic moments in interaction. In T. Stivers, L. Mondada & J. Steensig (Eds.), *The morality of knowledge in conversation* (pp. 159-183). Cambridge: Cambridge University Press.
- Heritage, J. (2012). Epistemics in action: Action formation and territories of knowledge. *Research on Language and Social Interaction*, 45(1), 1-29.
- Heritage, J. (2013). Epistemics in conversation. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 370-394). West Sussex: Blackwell.
- Heritage, J., & Atkinson, J. M. (1984). Introduction. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 1-15). Cambridge: Cambridge University Press.
- Heritage, J., & Clayman, S. (2010). *Talk in action: Interactions, identities and institutions*. Malden, MA: Wiley-Blackwell.
- Heritage, M., & Heritage, J. (2013). Teacher questioning: The epicenter of instruction and assessment. *Applied Measurement in Education*, 26(3), 176–190.
- Heritage, J., & Raymond, G. (2005). The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction. *Social Psychology Quarterly*, 68(1), 15-38.
- Heritage, J., & Raymond, G. (2012). Navigating epistemic landscapes: Acquiescence, agency and resistance in responses to polar questions. In J. P. De Ruiter (Ed.), *Questions: Formal, functional and interactional perspectives* (pp. 179-192). Cambridge: Cambridge University Press.
- Heritage, J., & Watson, D. R. (1979). Formulations as conversational objects. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 123-162). New York: Irvington Press.
- Hughes, A. (1989). *Testing for language teachers*. New York: Cambridge University Press.
- Hutchby, I., & Wooffitt, R. (1998). *Conversation analysis: Principles, practices and applications*. Oxford: Polity Press.
- Hutchby, I., & Wooffitt, R. (2008). *Conversation analysis* (2nd ed.). Cambridge: Polity Press.

- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402.
- Izard, J. (2005). Overview of test construction. In K. N. Ross (Ed.), *Quantitative research methods in educational planning*. Paris: UNESCO International Institute for Educational Planning.
- Jacknick, C. (2013). “Cause the textbook says...”: Laughter and student challenges in the ESL classroom. In P. Glenn & E. Holt (Eds.), *Studies of laughter in interaction* (pp. 185-200). London: Bloomsbury Academic.
- Jefferson, G. (1972). Side sequences. In D. Sudnow (Ed.), *Studies in social interaction* (pp. 294-338). New York: Free Press.
- Johansson, M. (2002). Sequential positioning of represented discourse in institutional media interaction. In A. Fetzer & C. Meierkord (Eds.), *Rethinking Sequentiality: Linguistics meets conversational interaction* (pp. 249-272). Amsterdam: Benjamins.
- Johnson, K. (2009). *Second language teacher education: A sociocultural perspective*. New York: Routledge.
- Kangasharju, H. (2002). Alignment in disagreement: Forming oppositional alliances in committee meetings. *Journal of Pragmatics*, 34(10-11), 1447-1471.
- Kasper, G., & Ross, S. J. (2007). Multiple questions in oral proficiency interviews. *Journal of Pragmatics*, 39(11), 2045-2070.
- Kasper, G., & Wagner, J. (2014). Conversation analysis in applied linguistics. *Annual Review of Applied Linguistics*, 34, 171-212.
- Kehoe, J. (1995). Writing multiple-choice test items. *Practical Assessment, Research & Evaluation*, 4(9), 1-3.
- Kelly, M., Grenfell, M., Allan, R., Kriza, C., & McEvoy, W. (2004). European profile for language teacher education: A frame of reference. *A Report to the European Commission Directorate General for Education and Culture*. University of Southampton.

- Kim, H. (2002). The form and function of next-turn repetition in English conversation. *Language Research* 38(1), 51-81.
- Kim, J., Chi, Y., Huensch, A., Jun, H., Li, H., & Roullion, V. (2010). A Case study on an item writing process: Use of test specifications, nature of group dynamics, and individual item writers' characteristics. *Language Assessment Quarterly*, 7(2), 160–174.
- Kitzinger, C. (2013). Repair. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 229-256). West Sussex: Blackwell.
- Kleinsasser, R. C. (2005). Transforming a postgraduate level assessment course: A second language teacher educator's narrative. *Prospect*, 20(3), 77–102.
- Koester, A. J. (2002). The performance of speech acts in workplace conversations and the teaching of communicative functions. *System*, 30(2), 167-184.
- Koester, A. (2006). *Investigating workplace discourse*. London: Routledge.
- Kormos, J. (1999). Simulating conversations in oral-proficiency assessment: a conversation analysis of role plays and non-scripted interviews in language exams. *Language testing*, 16(2), 163-188.
- Koshik, I. (2002). Designedly incomplete utterances: A pedagogical practice for eliciting knowledge displays in error correction sequences. *Research on Language and Social Interaction*, 35(3), 277-309.
- Kotthoff, H. (1993). Disagreement and concession in disputes: On the context sensitivity of preference structures. *Language in Society*, 22(2), 193-216.
- Lazaraton, A., & Davis, L. (2008). A microanalytic perspective on discourse, proficiency, and identity in paired oral assessment. *Language Assessment Quarterly*, 5(4), 313-335.
- Lazaraton, A., & Taylor, L. (2007). Qualitative research methods in language test development and validation. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner & C. Doe (Eds.), *Language testing reconsidered* (pp. 113-130). Ottawa: University of Ottawa Press.

- Lerner, G. H. (1996a). On the 'semi-permeable' character of grammatical units in conversation: Conditional entry into the turn space of another speaker. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 238-276). Cambridge: Cambridge University Press.
- Lerner, G. H. (1996b). Finding "face" in the preference structures of talk-in-interaction. *Social Psychology Quarterly*, 59(4), 303-321.
- Lerner, G. H. (2002). Turn-sharing: The choral co-production of talk-in-interaction. In C. E. Ford, B. A. Fox & S. A. Thompson (Eds.), *The language of turn and sequence* (pp. 225-256). Oxford, UK: Oxford University Press.
- Lerner, G. H. (2004). Collaborative turn sequences. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 225-256). Amsterdam and Philadelphia: John Benjamins.
- Lerner, G. H., & Kitzinger, C. (2010). Repair prefacing in the organization of same-turn self-repair. Paper presented at the International Conference on Conversation Analysis, Mannheim, Germany.
- Leung, C. (2009). Second language teacher professionalism. In A. Burns & J.C. Richards (Eds.), *The Cambridge guide to second language teacher education* (pp. 49-58). New York: Cambridge University Press.
- Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 103-130). West Sussex: Blackwell.
- Leyland, C., Greer, T., & Rettig-Miki, E. (2016). Dropping the devil's advocate: One novice language tester's shifting interactional practices across a series of speaking tests. *Classroom Discourse*, 7(1), 85-107.
- Liddicoat, A. J. (2007). *An introduction to conversation analysis*. New York: Continuum.
- Lindström, A., & Mondada, L. (2009). Assessments in social interaction: Introduction to the special issue. *Research on Language and Social Interaction*, 42(4), 299-308.
- Linn, R. L. (2006). The standards for educational and psychological testing: Guidance in test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 27-38). New Jersey: Lawrence Erlbaum.



- Luk, J. (2010). Talking to score: Impression management in L2 oral assessment and the co-construction of a test discourse genre. *Language Assessment Quarterly*, 7(1), 25-53.
- Lumley, T., & Brown, A. (2005). Research methods in language testing. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 833-856). New Jersey: Lawrence Erlbaum.
- Lynch, B. K., & Davidson, F. (1994). Criterion-referenced language test development: Linking curricula, teachers, and tests. *TESOL Quarterly*, 28(4), 727-743.
- MacDougall, C., & Baum, F. (1997). The devil's advocate: A strategy to avoid groupthink and stimulate discussion in focus groups. *Qualitative Health Research*, 7(4), 532-541.
- Madraso, J. (1993). Proofreading: The skill we've neglected to teach. *English Journal*, 82(2), 32-41.
- Markee, N. (2000). *Conversation analysis*. Mahwah, NJ: Lawrence Erlbaum.
- Markee, N. (2005). Conversation analysis for second language acquisition. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 355-374). New Jersey: Lawrence Erlbaum.
- Markee, N. (2007). Conversation analysis: Issues and problems. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 1017-1032). New York: Springer.
- Mathew, R. (2004). Stakeholder involvement in language assessment: Does it improve ethicality?. *Language Assessment Quarterly*, 1(2-3), 123-135.
- Maynard, D. W. (2013). Everyone and no one to turn to: Intellectual roots and contexts for conversation analysis. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 11-31). West Sussex: Blackwell.
- McNamara, T. (2004). Language testing. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 763-783). Oxford: Blackwell Publishing.
- McNamara, T. (2006). Assessment of second language proficiency. In K. Brown (Ed.), *Encyclopaedia of language and linguistics* (Vol. 1, pp. 546-553). Oxford: Elsevier.

- McNamara, T. (2007). Language testing: A question of context. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner & C. Doe (Eds.), *Language testing reconsidered* (pp. 131-137). Ottawa: University of Ottawa Press.
- McNamara, T., & Roever, C. (2006). *Language testing: The social dimension* (Vol. 1). Malden, MA: Blackwell.
- McQuade, R., Wiggins, S., Ventura-Medina, E., & Anderson, T. (2018). Knowledge disagreement formulations in problem-based learning tutorials: Balancing pedagogical demands with 'saving face'. *Classroom Discourse*, 9(3), 227-243.
- Milanovic, M. (2002). *Language examining and test development*. Strasbourg: Council of Europe Language Policy Division.
- Mislevy, R.J., Steinberg, L.S., & Almond, R.G. (2002). Design and analysis in task-based language assessment. *Language Testing*, 19(4), 477-496.
- Mondada, L. (2013). Conversation analysis and institutional interaction. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1005-1011). Oxford: Wiley-Blackwell.
- Myers, G. (1998). Displaying opinions: Topics and disagreement in focus groups. *Language in Society*, 27(1), 85-111.
- Nakatsuhara, F. (2008). Inter-interviewer variation in oral interview tests. *ELT Journal*, 62(3), 266-275.
- Nakatsuhara, F. (2013). *The co-construction of conversation in group oral tests*. Frankfurt: Peter Lang.
- Nissi, R. (2015). From entry proposals to a joint statement: Practices of shared text production in multiparty meeting interaction. *Journal of Pragmatics*, 79, 1-21.
- Nissi, R. (2016). Spelling out consequences: Conditional constructions as a means to resist proposals in organisational planning process. *Discourse Studies*, 18(3), 311-329.
- Ockey, G. J., Koyama, D., & Setoguchi, E. (2013). Stakeholder input and test design: A case study on changing the interlocutor familiarity facet of the group oral discussion test. *Language Assessment Quarterly*, 10(3), 292-308.

- Oğuz, Ş. (2003). State university preparatory class EFL instructors' attitudes towards assessment methods used at their institutions and portfolios as a method of alternative assessment. Unpublished Master's Thesis, Bilkent University.
- O'Loughlin, K. (2006). Learning about second language assessment: Insights from a postgraduate student on-line subject forum. *University of Sydney Papers in TESOL*, 1, 71–85.
- Peercy, M. M., Martin-Beltrán, M., Silverman, R. D., & Daniel, S. (2015). Curricular design and implementation as a site of teacher expertise and learning. *Teachers and Teaching: Theory and Practice*, 21(7), 867-893.
- Pekarek Doehler, S., & Petitjean, C. (2017). Introduction: Interactional competences in institutional settings- Young people between school and work. In S. Pekarek Doehler, A. Bangerter, G. De Weck, L. Filliettaz, E. González-Martínez & C. Petitjean (Eds.), *Interactional competences in institutional settings: From school to the workplace* (pp. 1-26). Cham: Palgrave Macmillan.
- Pomerantz, A. (1980). Telling my side: “Limited access” as a “fishing device”. *Sociological Inquiry*, 50(3-4), 186-198.
- Pomerantz, A. (1984a). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57-101). Cambridge: Cambridge University Press.
- Pomerantz, A. (1984b). Pursuing a response. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 152-163). Cambridge: Cambridge University Press.
- Pomerantz, A. (1990). Conversation analytic claims. *Communications Monographs*, 57(3), 231-235.
- Pomerantz, A., & Heritage, J. (2013). Preference. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 210-228). West Sussex: Blackwell.
- Psathas, G. (1995). *Conversation analysis: the study of talk-in-interaction* (Vol. 35). Thousand Oaks, CA: Sage.

- Raymond, G. (2003). Grammar and social organization: Yes/no interrogatives and the structure of responding. *American Sociological Review*, 68(6), 939-967.
- Robinson, J. D. (2013). Overall Structural Organization. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 257-280). West Sussex: Blackwell.
- Rossano, F. (2013). Gaze in conversation. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 308-329). West Sussex: Blackwell.
- Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button & J. R. E. Lee (Eds.), *Talk and social organisation* (pp. 54-69). Clevedon, UK: Multilingual Matters.
- Sacks, H. (1992). *Lectures on conversation* (Vols. 1 & 2). Oxford, UK: Blackwell.
- Sacks, H., Schegloff, E. A., & Jefferson, G. A. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50(4), 696-735.
- Sandlund, E., & Sundqvist, P. (2011). Managing task-related trouble in L2 oral proficiency tests: Contrasting interaction data and rater assessment. *Novitas-ROYAL (Research on Youth and Language)*, 5(1), 91-120.
- Sandlund, E., & Sundqvist, P. (2013). Diverging task orientations in L2 oral proficiency tests- a conversation analytic approach to participant understandings of pre-set discussion tasks. *Nordic Journal of Modern Language Methodology*, 2(1), 1-21.
- Sandlund, E., & Sundqvist, P. (2019). Doing versus assessing interactional competence. In M. R. Salaberry & S. Kunitz (Eds.), *Teaching and Testing L2 Interactional Competence: Bridging Theory and Practice* (pp. 357-396). New York: Routledge.
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309-327.
- Schegloff, E. A. (1992a). On talk and its institutional occasions. In P. Drew & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 101-134). Cambridge: Cambridge University Press.

- Schegloff, E. A. (1992b). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 97(5), 1295-1345.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge: Cambridge University Press.
- Schegloff, E. A., Koshik, I., Jacoby, S., & Olsher, D. (2002). Conversation analysis and applied linguistics. *Annual Review of Applied Linguistics*, 22, 3-31.
- Schoonen, R. (2011). How language ability is assessed. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 701-716). New York/London: Routledge.
- Seedhouse, P. (2004). Conversation analysis methodology. *Language Learning*, 54(S1), 1-54.
- Seedhouse, P. (2005). Conversation analysis as research methodology. In K. Richards & P. Seedhouse (Eds.), *Applying conversation analysis* (pp. 251-266). New York: Palgrave.
- Seedhouse, P. (2008). Learning to talk the talk: Conversation analysis as a tool for induction of trainee teachers. In S. Garton & K. Richards (Eds.), *Professional encounters in TESOL* (pp. 42-57). Basingstoke, UK: Palgrave MacMillan.
- Seedhouse, P. (2011). Conversation analytic research into language teaching and learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 345-363). New York/London: Routledge.
- Seedhouse, P. (2012). What kind of interaction receives high and low ratings in Oral Proficiency Interviews?. *English Profile Journal*, 3(1), 1-24.
- Seedhouse, P. (2013). Oral proficiency interviews as varieties of interaction. In S. J. Ross & G. Kasper (Eds.), *Assessing second language pragmatics* (pp. 199-219). London: Palgrave Macmillan.
- Seedhouse, P., Harris, A., Naeb, R., & Üstünel, E. (2014). Relationship between speaking features and band descriptors: A mixed methods study. *IELTS Research Reports Online Series*, 2, 1-30.

- Sert, O. (2015). *Social interaction and L2 classroom discourse*. Edinburgh: Edinburgh University Press.
- Sert, O. (2019). The interplay between collaborative turn sequences and active listenership: Implications for the development of L2 interactional competence. In M. R. Salaberry & S. Kunitz (Eds.), *Teaching and Testing L2 Interactional Competence: Bridging Theory and Practice* (pp. 142-166). New York: Routledge.
- Sert, O., & Seedhouse, P. (2011). Introduction: Conversation analysis in applied linguistics. *Novitas-ROYAL (Research on Youth and Language)*, 5(1), 1-14.
- Sevgi-Sole, E., & Ünalı, A. (2020). Rater negotiation scheme: How writing raters resolve score discrepancies. *Assessing Writing*, 45, 1-12.
- Sharma, B. K. (2012). Conceding in disagreements during small group interactions in academic writing class. *Classroom Discourse*, 3(1), 4-28.
- Shin, D. (2012). Item writing and writers. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 237-248). London/New York: Routledge.
- Shohamy, E. (1997). Second language assessment. In G. R. Tucker & D. Corson (Eds.), *Encyclopedia of language and education* (Vol. 4, pp. 141-149). Dordrecht: Springer.
- Shohamy, E. (2001). Democratic assessment as an alternative. *Language Testing*, 18(4), 373-391.
- Shohamy, E. (2007). The power of language tests, the power of the English language and the role of ELT. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 521-532). New York: Springer.
- Sidnell, J. (2010). *Conversation Analysis: An Introduction*. London: Wiley-Blackwell.
- Sidnell, J. (2013). Basic conversation analytic methods. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 77-99). West Sussex: Blackwell.
- Smithson, J., & Díaz, F. (1996). Arguing for a collective voice: Collaborative strategies in problem-oriented conversation. *Text & Talk*, 16(2), 251-268.

- So, Y. (2014). Are teacher perspectives useful? Incorporating EFL teacher feedback in the development of a large-scale international English test. *Language Assessment Quarterly*, 11(3), 283-303.
- Spaan, M. (2007). Evolution of a test item. *Language Assessment Quarterly*, 4(3), 279-293.
- Stevanovic, M. (2012). Establishing joint decisions in a dyad. *Discourse Studies*, 14(6), 779-803.
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238-245.
- Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18(1), 23-27.
- Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5-15.
- Stivers, T. (2005). Modified repeats: One method for asserting primary rights from second position. *Research on Language and Social Interaction*, 38(2), 131-158.
- Stivers, T. (2011). Morality and question design: 'Of course' as contesting a presupposition of askability. In T. Stivers, L. Mondada & J. Steensig (Eds.), *The morality of knowledge in conversation* (pp. 82-106). Cambridge: Cambridge University Press.
- Stivers, T. (2013). Sequence organization. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 191-209). West Sussex: Blackwell.
- Suryaningsih, H. (2014). Students' perceptions of international English language testing system (IELTS) and test of English as a foreign language (TOEFL) tests. MA Thesis, Indiana University of Pennsylvania.
- Szczepek, B. (2000). Functional aspects of collaborative productions in English conversation. *Interaction and Linguistic Structures*, 21, 1-36.
- Şallı-Çopur, D. (2008). Teacher effectiveness in initial years of service: A case study on the graduates of METU Foreign Language Education Program. Unpublished doctoral dissertation, Middle East Technical University, Ankara.

- Tan, M., & Turner, C. E. (2015). The impact of communication and collaboration between test developers and teachers on a high-stakes ESL exam: Aligning external assessment and classroom practices. *Language Assessment Quarterly*, 12(1), 29-49.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21-26.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412.
- Ten Have, P. (2007). *Doing conversation analysis: A practical guide* (2nd ed.). London: Sage.
- Van Kruiningen, J.F. (2013). Educational design as conversation: A conversation analytical perspective on teacher dialogue. *Teaching and Teacher Education*, 29, 110-121.
- Verkuyten, M. (2000). School marks and teachers' accountability to colleagues. *Discourse Studies*, 2(4), 452-472.
- Walsh, S. (2003). Developing interactional awareness in the second language classroom through teacher self-evaluation. *Language Awareness*, 12(2), 124-142.
- Walsh, S. (2011). *Exploring classroom discourse: Language in action*. Oxon: Routledge.
- Waring, H. Z. (2001). Balancing the competing interests in seminar discussion: Peer referencing and asserting vulnerability. *Issues in Applied Linguistics*, 12(1), 29-50.
- Waring, H. Z. (2017). Going general as a resource for doing advising in post-observation conferences in teacher training. *Journal of Pragmatics*, 110, 20-33.
- Wasson, C. (2000). Caution and consensus in American business meetings. *Pragmatics*, 10(4), 457-481.
- Wasson, C. (2016). Integrating conversation analysis and issue framing to illuminate collaborative decision-making activities. *Discourse & Communication*, 10(4), 378-411.



- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Oxford: Palgrave Macmillan.
- Welch, C. (2006). Item and prompt development in performance testing. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 91-114). New Jersey: Lawrence Erlbaum.
- Wolf, M. K., Farnsworth, T., & Herman, J. (2008). Validity issues in assessing English language learners' language proficiency. *Educational Assessment*, 13(2-3), 80-107.
- Wong, J., & Waring, H. Z. (2010). *Conversation analysis and second language pedagogy: A guide for ESL/EFL teachers*. New York: Routledge.
- Xerri, D., & Briffa, P. V. (2018). *Teacher involvement in high-stakes language testing*. Cham: Springer.
- You, H. J. (2015). Reference to shared past events and memories. *Journal of Pragmatics*, 87, 238-250.
- Young, R. F., & He, A. W. (1998). *Talking and testing: Discourse approaches to the assessment of oral proficiency* (Vol. 14). Amsterdam: John Benjamins.
- Zimmerman, D. H., & Boden, D. (1991). Structure-in-action: An introduction. In D. Boden & D. H. Zimmerman (Eds.), *Talk and social structure: Studies in ethnomethodology and conversation analysis* (pp. 3-21). Cambridge: Polity.

## APPENDICES

### A. BACKGROUND QUESTIONNAIRE

**For teachers:**

#### A. DEMOGRAPHIC INFORMATION

1. Name-Surname: .....

2. Age: .....

3. Gender: ☐Female ☐Male

4. Marital Status: ☐Married ☐Single

5. What is your educational background?

B.A. University/Department: ...../.....

M.A. ☐None ☐Completed ☐In progress

University/Department: ...../.....

Ph.D. ☐None ☐Completed ☐In progress

University/Department: ...../.....

6. What is your job experience? Please fill in the table below.

| Institution name | Responsibilit(ies)<br>(e.g., teaching, testing,<br>curriculum designer, etc.) | State or Private? | Duration<br>(Years/Month) |
|------------------|-------------------------------------------------------------------------------|-------------------|---------------------------|
| a.               |                                                                               |                   |                           |
| b.               |                                                                               |                   |                           |
| c.               |                                                                               |                   |                           |
| d.               |                                                                               |                   |                           |

7. How many years/months have you been teaching English? .....

8. How many hours of teaching do you generally have a week?.....

9. What are your areas of interest?.....

10. Have you ever worked at a testing unit? (i.e., Do you have any testing experience?)

☐Yes ☐No

If yes,

a) At which institution(s) did you do testing? / For how long (years/months)?

...../.....  
...../.....

b) What was it like to work at the testing unit?

.....  
.....

c) What helped you most to improve your testing skills at the testing unit?

.....

d) How have you benefitted from this testing experience?

.....

11. Are you one of the 'reviewers' of exams at your institution? If yes, why?

.....

## B. VIEWS ON ITEM REVIEWING PRACTICE IN TEST CONSTRUCTION

1. What do you think about the item reviewing practice at testing units?

.....

2. What makes an item reviewing session in language test construction more effective?

.....

**For testers:**

## A. DEMOGRAPHIC INFORMATION

1. Name-Surname: .....

2. Age: .....

3. Gender: ☐Female ☐Male

4. Marital Status: ☐Married ☐Single

5. What is your educational background?

B.A. University/Department: ...../.....

M.A. ☐None ☐Completed ☐In progress

University/Department: ...../.....

Ph.D. ☐None ☐Completed ☐In progress

University/Department: ...../.....

6. What is your job experience? Please fill in the table below.

| Institution name | Responsibilit(ies)<br>(e.g., teaching, testing,<br>curriculum designer, etc.) | State or Private? | Duration<br>(Years/Month) |
|------------------|-------------------------------------------------------------------------------|-------------------|---------------------------|
| a.               |                                                                               |                   |                           |
| b.               |                                                                               |                   |                           |
| c.               |                                                                               |                   |                           |
| d.               |                                                                               |                   |                           |

7. How many years/months have you been teaching/did you teach English?.....

8. What are your areas of interest?.....

9. As an examiner,

- a) How many hours of teaching do you generally have a week? .....
- b) How many years/months have you been working at the testing unit?.....
- c) Have you worked at a testing unit before in a different institution?

☐Yes ☐No

If yes,

At which institution(s)?                      /                      For how long (years/months)?  
...../.....  
...../.....

- d) What is it like to work at the testing unit?  
.....  
.....
- e) What helps you most to improve your testing skills at the testing unit?  
.....  
.....
- f) How do you benefit from this testing experience?  
.....  
.....

## **B. VIEWS ON ITEM REVIEWING PRACTICE IN TEST CONSTRUCTION**

- 1. What do you think about the item reviewing practice at testing units?  
.....  
.....
- 2. What makes an item reviewing session in language test construction more effective?  
.....  
.....
- 3. What are your criteria in choosing your 'reviewers'?  
.....  
.....

## B. JEFFERSON TRANSCRIPTION CONVENTIONS

Adapted from Hutchby and Wooffitt (2008)

|              |                                                                                                                                                                                                                             |
|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (1.8)        | Numbers enclosed in parentheses indicate a pause. The number represents the number of seconds of duration of the pause, to one decimal place. A pause of less than 0.2 seconds is marked by (.)                             |
| [ ]          | Brackets around portions of utterances show that those portions overlap with a portion of another speaker's utterance.                                                                                                      |
| =            | An equal sign is used to show that there is no time lapse between the portions connected by the equal signs. This is used where a second speaker begins their utterance just at the moment when the first speaker finishes. |
| ::           | A colon after a vowel or a word is used to show that the sound is extended. The number of colons shows the length of the extension.                                                                                         |
| (hm, hh)     | These are onomatopoeic representations of the audible exhalation of air)                                                                                                                                                    |
| .hh          | This indicates an audible inhalation of air, for example, as a gasp. The more h's, the longer the in-breath.                                                                                                                |
| ?            | A question mark indicates that there is slightly rising intonation.                                                                                                                                                         |
| .            | A period indicates that there is slightly falling intonation.                                                                                                                                                               |
| ,            | A comma indicates a continuation of tone.                                                                                                                                                                                   |
| -            | A dash indicates an abrupt cut off, where the speaker stopped speaking suddenly.                                                                                                                                            |
| ↑↓           | Up or down arrows are used to indicate that there is sharply rising or falling intonation. The arrow is placed just before the syllable in which the change in intonation occurs.                                           |
| <u>Under</u> | Underlines indicate speaker emphasis on the underlined portion of the word.                                                                                                                                                 |
| CAPS         | Capital letters indicate that the speaker spoke the capitalized portion of the utterance at a higher volume than the speaker's normal volume.                                                                               |
| °            | This indicates an utterance that is much softer than the normal speech of the speaker. This symbol will appear at the beginning and at the end of the utterance in question.                                                |
| ><, <>       | 'Greater than' and 'less than' signs indicate that the talk they surround was noticeably faster, or slower than the surrounding talk.                                                                                       |

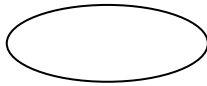
|                |                                                                                                                                                                                                                                                 |
|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (would)        | When a word appears in parentheses, it indicates that the transcriber has guessed as to what was said, because it was indecipherable on the tape. If the transcriber was unable to guess what was said, nothing appears within the parentheses. |
| £C'mon£        | Sterling signs are used to indicate a smiley or jokey voice.                                                                                                                                                                                    |
| +              | marks the onset of a non-verbal action (e.g. shift of gaze, pointing)                                                                                                                                                                           |
| <i>italics</i> | English translation                                                                                                                                                                                                                             |

### Other Symbols

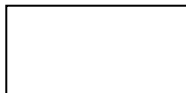
|           |                                                                                           |
|-----------|-------------------------------------------------------------------------------------------|
| no#...#no | The onset and offset point of a non-verbal activity accompanying the talk (Balaman, 2016) |
|-----------|-------------------------------------------------------------------------------------------|

### C. MEANINGS OF FLOWCHART SYMBOLS USED IN FIGURES

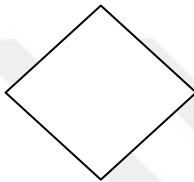
Adapted from Goetsch and Davis (2014)



An oval is used to indicate the start or end of a process.



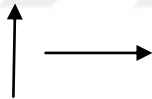
A rectangle is used to indicate a process step, an activity, or an operation. It may also represent an entire sub-process.



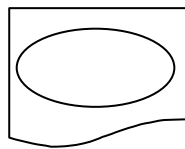
A diamond is used to show a decision point. This is usually a “yes” or “no” function but could also represent an “if/then” statement. The process flow branches off in two or more directions, depending on the decision.



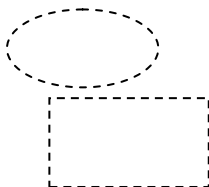
This symbol represents documentation, report or paperwork produced or required by the process.



Arrows indicate the sequential path and direction of flow in the process. In this study, they show the sequential organization of the interactions examined.



More than one symbol is used when two activities are performed at the same time in the process. For example, this symbol used in Figure 4 shows that the step involves paperwork as the decided change is marked on the exam paper, which is at the same time the end of the interactional process.



Dashed lines around symbols are used to indicate an optional activity or step in the process.

## D. APPROVAL OF METU HUMAN SUBJECTS ETHICS COMMITTEE

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ  
APPLIED ETHICS RESEARCH CENTER



ORTA DOĞU TEKNİK ÜNİVERSİTESİ  
MIDDLE EAST TECHNICAL UNIVERSITY

Sayı: 2016-EGT-003  
SAYI: 2016-EGT-003  
ÇANKAYA ANKARA/TURKEY  
T: +90 312 210 22 91  
F: +90 312 210 79 59  
www.ueam.metu.edu.tr

27 OCAK 2016

Gönderen: Prof. Dr. Çiler HATİPOĞLU

Eğitim Fakültesi

Gönderen: Prof. Dr. Canan SÜMER

İnsan Araştırmaları Komisyonu Başkanı

İlgi: Etik Onayı

Sayın Prof. Dr. Çiler HATİPOĞLU 'nun danışmanlığını yaptığı Hümeyra CAN'ın "An Interactive Perspective to Language Test Construction: Implications for Language Assessment Literacy (LAL) " başlıklı araştırması İnsan Araştırmaları Komisyonu tarafından uygun görülerek gerekli onay 2016-EGT-003 protokol numarası ile 15.02.2016-31.12.2017 tarihleri arasında geçerli olmak üzere verilmiştir.

Bilgilerinize saygılarımla sunarım.

Prof. Dr. Canan SÜMER

Uygulamalı Etik Araştırma Merkezi

İnsan Araştırmaları Komisyonu Başkanı

Prof. Dr. Meliha ALTUNİŞİK

İnsan Araştırmaları Komisyonu

Üyesi

(İZİNÜ)

Prof. Dr. Aydan BALAMİR

İnsan Araştırmaları Komisyonu

Üyesi

Prof. Dr. Mehmet UTKU

İnsan Araştırmaları Komisyonu

Üyesi

Prof. Dr. Ayhan SOL

İnsan Araştırmaları Komisyonu

Üyesi



## E. INFORMED CONSENT FORM

This study aims to bring an interactional perspective to language test construction by presenting and describing the micro-details of the recurring patterns and interactional features that emerge in interactions between teachers and examiners in item reviewing sessions where they critically review and co-construct language tests collaboratively. In doing this, video and audio recordings of syllabus-based classroom language test item reviewing sessions taking place among classroom language teachers and an examiner will be collected and transcribed for analysis. In addition, if required, after the initial analysis of the data, stimulated recall sessions will be held with the participants with regard to the interactional patterns that have emerged. In these sessions, the participants will be asked to watch the relevant videos and share their interpretations, explanation, aim, thoughts and feelings regarding the events. This study is expected to contribute to professional development and second language teacher education programs and more specifically, to teachers' language assessment literacy, their test item reviewing skill and interactional competence in a goal-oriented institutional activity with stakeholders.

Participation in the study must be on a voluntary basis. No personal identification information will be asked during the study and any information obtained by chance will be kept confidential. For example, when transcribing the recorded data, pseudonyms will be used instead of the actual names. The recordings will be kept strictly confidential and will be analysed and evaluated only by the researchers; the obtained data will be used for scientific purposes only. During participation, if you feel uncomfortable for any reason, you are free to stop the video/audio recording at any time.

If you have any questions regarding the study, please do not hesitate to contact the researchers. We would like to thank you in advance for your participation in this study. For further information about the study, contact information of the researchers is given below:

### Researcher

Hümeýra Can  
Middle East Technical University  
PhD Candidate  
E-mail: [hmeýra.can@gmail.com](mailto:hmeýra.can@gmail.com)

### Advisor

Assoc. Prof. Dr. Çiler Hatipoğlu  
Middle East Technical University  
Foreign Language Education  
Phone: 0 312 210 40 75  
E-mail: [ciler.hatipoglu@gmail.com](mailto:ciler.hatipoglu@gmail.com)

*I am participating in this study totally on my own will and am aware that I can stop the video/audio recording at any time I want/ I give my consent for the use of the information I provide for scientific purposes.* (Please return this form to the data collector after you have filled it in and signed it).

**Date:**

### Participant

Name Surname:  
Phone:  
E-mail:  
Signature:

## F. CURRICULUM VITAE

### PERSONAL INFORMATION

Surname, Name: Can, Hümeýra  
Nationality: Turkish (TC)  
Date and Place of Birth: 17 July 1985, Germany  
Marital Status: Single  
E-mail: hmeýra.can@gmail.com

### EDUCATION

| Degree      | Institution                               | Year of Graduation |
|-------------|-------------------------------------------|--------------------|
| MS          | METU English Language Teaching            | 2011               |
| BS          | METU English Language Teaching            | 2008               |
| High School | Roxburgh College,<br>Melbourne, Australia | 2003               |

### WORK EXPERIENCE

| Year      | Place                                                | Enrollment         |
|-----------|------------------------------------------------------|--------------------|
| 2008-2016 | Hacettepe University, School of<br>Foreign Languages | English Instructor |
| 2008      | Ufuk University, English<br>Preparatory School       | English Instructor |

### FOREIGN LANGUAGES

Advanced English

### PUBLICATIONS

- Can, H. (2017). "We want to great husband or wife": An analysis of gerund and infinitive errors made by Turkish university students in EFL written discourse. In D. Koban Koç (Ed.), *Writing: Suggestions for improvement in ESL and EFL Contexts* (pp. 92-128). Berlin: Dağyeli Verlag. ISBN: 9783935597265.
- Can, H. (2011). Türkçe tebrikler: E-posta grup mesajlarına bir bakış (Turkish congratulations: A look at group e-mail messages). *Proceedings of the 24<sup>th</sup> National Linguistics Conference*, Middle East Technical University, Ankara, pp. 121-131.
- Hatipoğlu, Ç., & Can, H. (2011). The intercultural approach to teaching Turkish as a foreign language: A closer look at e-mail congratulation messages. *Proceedings of the 3<sup>rd</sup> International Turkish Language Education Symposium*, Dokuz Eylül University, İzmir, pp. 185-197.
- Can, H., & Can, N. (2010). The inner self desires a friendly chat: Chat metaphors in Turkish and English. *Metaphor and Symbol*, 25(1), 34-55.
- Ruhi, Ş., Eröz-Tuğa, B., Hatipoğlu, Ç., Işık-Güler, H., Acar, G. C., Eryılmaz, K., Can, H., Karakaş, Ö., & Çokal Karadaş, D. (2010). Sustaining a corpus for spoken Turkish discourse: Accessibility and corpus management issues. *Proceedings of the Language*

*Resources: From Storyboard to Sustainability and LR Lifecycle Management Workshop, Malta, pp. 44-48.*

Can, H. (2010). Refreshing our minds at the INGED-Hacettepe Day. *The INGED Newsletter, 1*, 25-28.

Can, H. (2009). Reflections from the 10<sup>th</sup> METU international ELT conference: “Growth”. *The INGED Newsletter, 3*, 47-48.

## PROJECTS

| Year      | Project Type & Title                                                                                                                                                                                          | Position   |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 2008-2011 | TUBITAK- EVRENA, 108K283<br>Development of “METU Spoken Turkish Corpus”                                                                                                                                       | Researcher |
| 2014-2015 | BAB Project 1714- Hacettepe University<br>014 D06 704 001<br>An Analysis and Development of Writing Skills of<br>English Language Learners Enrolled in the English<br>Preparatory School of Foreign Languages | Researcher |

## CONFERENCE PRESENTATIONS

Can, H., & Hatipoğlu, Ç. (July 2015). L2 Turkish learners’ conceptualization of the ‘polite’ speech act of congratulation in Turkish. *9<sup>th</sup> International Conference of Im/politeness*, National and Kapodistrian University of Athens, Athens, Greece.

Can, H. (July 2015). The speech act of criticism and (im)politeness in intercultural communication: The case of L2 Turkish. *9<sup>th</sup> International Conference of Im/politeness*, National and Kapodistrian University of Athens, Athens, Greece.

Can, H. (March 2015). Classroom interactive decisions beyond teacher lesson plans: A Case from an EFL classroom. *AAAL2015 Conference*, Toronto, Canada.

Can, H. (April 2014). Teacher cognition and language classrooms: Investigating what’s going on out there. *8<sup>th</sup> International IDEA Conference: Studies in English*, Muğla Sıtkı Koçman University, Muğla, Turkey.

Can, N., & Can, H. (June 2013). Exploring pre-service EFL teachers’ needs for oral communication skills. *13<sup>th</sup> International ELT Conference: Teachers Exploring Practice for Professional Learning*, Bilkent University, Ankara, Turkey.

Can, H. (October 2011). Data-driven learning approach: Using corpus in vocabulary teaching. *INGED 15<sup>th</sup> International ELT Conference: Taking it to the Limits*, Hacettepe University, Ankara, Turkey.

Hatipoğlu, Ç., & Can, H. (July 2011). A corpus approach to the cross-cultural study of the ‘polite’ speech act of congratulation in British English and Turkish. *6<sup>th</sup> International Symposium on Politeness*, Middle East Technical University, Ankara, Turkey.

Hatipoğlu, Ç., & Can, H. (July 2010). New data collection techniques, new materials and the intercultural approach to teaching Turkish as a second/foreign language. *III. International Turkish Language Education Symposium*, Dokuz Eylül University, İzmir, Turkey.

Can, H. (May 2010). Türkçe tebrikler: E-posta grup mesajlarına bir bakış (Turkish congratulations: A look at group e-mail messages). *24<sup>th</sup> National Linguistics Conference*, Middle East Technical University, Ankara, Turkey.

## G. TURKISH SUMMARY / TRKE ZET

### İNGİLİZCEYİ YABANCI DİL OLARAK ÖĞRETEN ÖĞRETMENLERİN DİL SINAV SORULARINI İNCELEME ETKİLEŞİMLERİNİN MİKRO ÇÖZÜMLEME YOLUYLA İNCELENMESİ

Bu çalışma Konuşma Çözümlemesi (KÇ) kullanarak müfredata dayalı İngilizce dil sınavlarının oluşturulmasına, daha belirgin olarak madde inceleme (Mİ) aşamasına etkileşimsel bir bakış açısı getirmektedir. Dil sınavı oluşturma modellerinde Mİ önemli aşamalardan biri olarak ele alınmasına ve bu aşamada meslektaşlarla olan etkileşimin önemi vurgulanmasına rağmen bu uygulamanın daha iyi anlaşılmasını sağlayacak Mİ etkileşimlerinin mikro-ayrıntıları çalışılmamıştır. Dolayısıyla, bu çalışmada Türkiye’deki bir devlet üniversitesinin İngilizce hazırlık okulunda gerçekleşen Mİ oturumlarının video kayıtlarından (25 saat) oluşan bir veritabanı kullanılarak, İngilizceyi yabancı dil olarak öğreten öğretmenlerin öğrencileri için hazırlanmış olan dil sınav sorularını sınavcı ile birlikte etkileşim yoluyla nasıl inceledikleri araştırılmıştır.

Türkiye’de üniversitelerdeki İngilizce hazırlık okullarının amacı öğretim dili İngilizce olan bölümlerdeki dersleri etkili bir şekilde takip edebilmeleri için üniversiteye yeni başlayan öğrencilerin İngilizcede yeterli bilgi ve becerilere sahip olmalarını sağlamaktır. Birçok hazırlık okulundaki uygulamalara benzer olarak, bu çalışma için verinin elde edildiği okulda kalabalık bir öğrenci ve öğretmen grubu mevcut olduğu için standartlığı sağlamak adına dil sınavları yerel olarak okulun sınav birimleri tarafından hazırlanmaktadır. Bu birimlerde ise hale hazırda görev yapan öğretmenler görevlendirilmektedir ve bu öğretmenler sorumlu oldukları dil seviyesindeki öğrenci gruplarının sınavlarını hazırlamakla yükümlüdürler. Bu süreçte, sınavcılar müfredatı ve sınıfta işlenen materyalleri temel alarak hazırladıkları sınavları derse giren öğretmenlerin (en az bir kişi) özellikle öğretim ve öğrenme ortamlarını göz önünde bulundurarak inceleyip değerlendirmelerini ve etkileşim içinde geribildirimde bulunmalarını istemektedirler. Bu çalışma ise sınavcı ve öğretmenler arasında geçen bu dil sınavı madde inceleme (Mİ) etkileşimlerine odaklanmıştır.

Söz konusu araştırma bağlamında, Mİ uygulaması sınav geliştirme sürecinin önemli bir parçasıdır. Bunun en belirgin nedeni Mİ etkileşimlerinin sınavların geçerlik ve güvenilirliğinin sağlanmasına; öğrenci, öğretmen ve sınavcı gibi ilgili kişileri olumsuz yönde etkileyebilecek olası problemlerin öngörülmesine yardımcı olmasıdır. Mİ uygulaması için genellikle hem öğretmenlikte hem de sınav hazırlamada tecrübesi olan ve yapıcı geribildirimde bulunabilen öğretmenler seçilmektedir. Sınav Mİ görevini yapan bu öğretmenler sınav sorularını bir öğrenci gibi cevaplamakta ve her birini eleştirel bir bakış açısıyla incelemektedirler. Öğretmenler sınav sorularını öğrencileri, sınıf ortamı, öğretim etkinlikleri, program içeriği, sınav geliştirme bilgi ve tecrübelerini göz önünde bulundurarak incelemekte; böylelikle öğretim ve sınav uygulamaları arasında köprü görevi görmektedirler. Öğretmenlerin sadece sorulardaki problemleri tespit edip açıklamaları değil, aynı zamanda problemleri sorular için önerilerde bulunup bunları sınavcılarla işbirliği içinde düzeltmeleri ve geliştirmeleri beklenmektedir.

Bu çalışmaya toplam 7 sınavcı (okulun sınav biriminde görevlendirilmiş öğretmenler) ve 25 öğretmen (Mİ işlemini yapanlar) katılmıştır. Mİ görevini yapan öğretmenlerin hepsi mesleklerinde tecrübeli, birçoğu ise sınav biriminde görev almış ve dolayısıyla sınav hazırlamada tecrübeli kişilerdir. Ayrıca bu öğretmenler sıklıkla Mİ oturumlarına davet edilen sınav inceleme görevinde kurumun tecrübeli öğretmenleri arasındadır. Bundaki en büyük faktör bu öğretmenlerin sınav birimindeki önceki tecrübeleri, işbirlikçi, dikkatli ve analitik yaklaşımlarıyla Mİ'ye olan olumlu tutumlarıdır. Katılımcılardan sınavcı grubuna bakıldığında ise bu öğretmenlerin öğretmenlik mesleğinde tecrübeli olmalarına rağmen Mİ görevini yerine getiren öğretmenlerden nispeten daha genç oldukları görülmektedir.

İngilizce dil sınavı Mİ oturumlarının video kayıtları alınırken araştırmacının katılımı ve hiçbir şekilde müdahalesi olmamıştır. Dolayısıyla, bu çalışmanın veritabanı doğal olarak gelişen etkileşimlerden oluşmaktadır. Video kayıtlarına başlamadan önce gerekli etik kurul izni alınmıştır. Ayrıca katılımcılara çalışmayla ilgili gerekli bilgilendirme yapılmış, gönüllü katılım formları dağıtılarak kendilerinin izinleri alınmıştır. Nitel çalışmaların bir gerekliliği olarak çalışmanın gerçekleştiği bağlam ve katılımcılarla ilgili ayrıntılı açıklama yapılabilmesi için video kayıtlarından sonra katılımcılara bir anket verilmiştir. Bu anket yoluyla katılımcıların demografik bilgileri alınmış, öğretmenlik ve sınav hazırlama tecrübeleri, kurumlarındaki Mİ uygulamasıyla ilgili görüşleri öğrenilmiştir. Kaydedilen oturumlar çoğunlukla iki (bir öğretmen-bir sınavcı) veya üç (iki öğretmen-bir sınavcı) kişilik

gruplar halinde gerçekleşmiştir. Mİ eylemi görev-odaklı olup küçük gruplardan oluştuğu için video kayıtları için bir kamera ve iki ses kaydedici kullanılarak gerekli alt yapı oluşturulmuş, etkileşimlerin çok-kipli (multi-modal) bir aktivite olarak kaydedilmesi sağlanmıştır. Toplamda çeşitli dil seviyelerinde 26 Mİ oturumu kaydedilmiştir. Veritabanındaki video kayıtlarında incelenen sınavlar ağırlıklı olarak dil kullanımı bölümü (gramer ve kelime), bunu takiben de sırasıyla okuma ve dinleme bölümlerinden oluşmaktadır. Bu sınavlarda çeşitli soru tiplerinin kullanılmasıyla birlikte en sık kullanılan soru tipleri çoktan-seçmeli, doğru/yanlış ve boşlukları-doldurma olmuştur. İncelenen sınavlar program içeriğine dayalı bir yaklaşım izlemiş, öğrencilerin dil gelişimini ve başarısını sınıfta işlenen materyaller kapsamında değerlendirmeyi amaçlamıştır.

Bu çalışmanın veri analizi Konuşma Çözümlemesi (KÇ) yöntemi kullanılarak yapılmıştır. Video kayıtlarını takiben elde edilen verinin çeviriyazıları oluşturularak Mİ etkileşimlerinin yapısını ve tekrar eden etkileşimsel eylemleri tespit edebilmek için “nedensiz bakış” (unmotivated looking) yoluyla verinin ön incelemesi yapılmıştır. Bu süreçte bazı kesitler HUMAN Araştırma Merkezinin veri oturumlarında sunulmuştur. Ön incelemeler Mİ etkileşimlerin daha çok problemleri odaklandığını göstermiştir. Bu durumlarda sınav maddelerindeki problemlerin konuşma içinde ve konuşma aracılığıyla nasıl ortaya çıktığı, Mİ grubunun bu problemleri etkileşimsel olarak nasıl tartıştığı ve bunlara nasıl karar verdiği, bunu yaparken ne tür sınav hazırlama ilkelerine yöneldikleri ve problemleri sınav maddelerini düzelterek geliştirirken nasıl bir yol izledikleri gözlemlenmiştir. Bunun üzerine, Mİ etkileşimlerinde sınav maddelerinin problemleştirildiği kesitlerden koleksiyon oluşturulmuştur. Ana koleksiyonda yapılan ön analiz ise problemleri tartışıldığı etkileşimlerin tekrarlayan üç aşamadan oluştuğunu göstermiştir; (1) problemi belirleme (PB), (2) değişiklik önerme (DÖ), ve (3) değişikliği gözden geçirme (DG). Bu genel etkileşimsel düzenin tespit edilmesi üzerine PB, DÖ ve DG aşamaları için ayrı ayrı alt koleksiyonlar oluşturulmuştur. Daha sonra bu alt koleksiyonlardan temsili kesitler belirlenmiş, bunların detaylı çeviriyazısı ve analizleri yapılmıştır. Analizleri desteklemek için seçilen kesitlerden dört tanesi veri oturumlarında sunulmuştur. Analizler KÇ yöntemi kullanılarak yapılmıştır.

KÇ yöntemi “insan sosyal hayatının temel ve belirleyici bir özelliği olarak konuşmayı açıklamayı, analiz etmeyi ve anlamayı hedeflemektedir” (Sidnell, 2010, s.1). Dolayısıyla, KÇ doğal gelişen etkileşimleri incelemeyi kapsayan doğası gereği nitel bir yöntemdir

(Hutchby ve Wooffitt, 1998; Markee, 2000). Etkileşimsel desenler veriden çıkarıldığı için ve araştırmacı veriye dışarıdan daha önce belirlenmiş kategori, kod veya herhangi bir kuramsal varsayım getirmedeği için KÇ veri analizine ampirik ve veri-güdümlü bir yaklaşım sunmaktadır (Seedhouse, 2004; Walsh, 2011). KÇ ilkelerine göre etkileşimi düzenleyen ve dolayısıyla günlük hayattaki sosyal eylemlerin temel yapısını oluşturan etkileşimsel kaynaklar mevcuttur: (1) söz sırası alma (turn-taking), (2) dizi düzeni (sequence organization), (3) onarım (repair), (4) söz sırası tasarımı (turn design) (Wong ve Waring, 2010). Etkileşim içindeki kişiler bu etkileşimsel kaynakları kullanarak öznelarasılığı (intersubjectivity) ve karşılıklı anlaşılabilirliği sağlarlar (Seedhouse, 2004). KÇ, katılımcıların konuşmayı üretmede ve anlamada kullandıkları etkileşimsel kaynaklara mikro-analitik bir yaklaşımla erişim sağlar ve böylelikle konuşmayı içeriden bir bakış açısıyla (emic perspective) çözümler (Gardner, 2004; Seedhouse, 2004). Diğer bir ifadeyle, KÇ, diziler içindeki söz sıralarının mikro ayrıntılarını açıklayarak herhangi bir varsayımda bulunmadan katılımcıların etkileşim içindeki yönelimlerini ve konuşmanın ayrıntılarında sundukları kanıtları ortaya çıkarır (Psathas, 1995; Seedhouse, 2004).

KÇ’de önemli gelişmelerden birisi de kurumsal etkileşimlere yönelik çalışmalardır (Drew ve Heritage, 1992; Heritage, 1997; Heritage ve Clayman, 2010). Bu çalışmalar kurumsal ortamlarda profesyonellerin günlük aktivitelerini gerçekleştirmede ve yönetmede dahil oldukları etkileşimleri incelemiştir. Kurumsal ortamlardaki konuşma hedef ve görev odaklı olup etkileşimin düzeni, katılımcıların etkileşimsel katkılarını sınırlayan kurumsal hedefe bağlıdır (Drew ve Heritage, 1992; Gardner, 2004; Heritage, 2005; Heritage ve Clayman, 2010; Seedhouse, 2011; Walsh, 2011). Bu ise yeğleme düzenini (preference organization) ve katılımcıların dizisel, söz sırası alma, onarım uygulamalarını günlük konuşmadan oldukça farklı yapmaktadır (Markee, 2007). Diğer bir ifadeyle, kurumsal bağlamlarda katılımcılar hareketlerini kuruma özel hedef ve kısıtlamalara göre şekillendirir; bu ise kurumsal etkileşime ayırt edici, özelleşmiş, sınırlı ve tekrar eden bir özellik kazandırmaktadır (Drew ve Heritage, 1992; Heritage ve Clayman, 2010; Zimmerman ve Boden, 1991). Ayrıca, KÇ’ye göre kurumun belli bir konuşma için bağlam sağlamasından ziyade bağlam konuşma içinde ve konuşma aracılığıyla katılımcılar tarafından oluşturulmakta ve katılımcılar görevlerini konuşma yoluyla kurumsal hedefler doğrultusunda yerine getirmektedirler (Drew & Heritage, 1992; Heritage, 1997; 2005). Etkileşimin kurumsal özelliğini incelemek için günlük konuşmaları çözümlemede kullanılan temel etkileşimsel kaynakların yanı sıra genel yapısal düzen (overall structural organization), sözcüksel seçim (lexical choice) ve

etkileşimsel asimetri biçimleri (forms of interactional asymmetry) gibi boyutların da ele alınması önerilmiştir (Drew ve Heritage, 1992). Bu çalışma da bir iş yeri çalışması olup bir yükseköğrenim kurumunda öğretmenler arasında geçen görev-odaklı etkileşimlerden kurumda yaygın bir uygulama olan dil sınavı Mİ etkileşimlerini inceleyerek öğretmenlerin bu uygulamaya dair farkındalığını artırmayı hedeflemektedir. Böylelikle bu çalışma literatürde dil sınavı geliştirmenin önemli bir aşaması olarak ele alınan Mİ görevine yönelik yeni bir anlayış geliştirerek dil değerlendirme ve dil sınavbilimi alanları için uygulamalar sunmakta, kurumsal KÇ'ye katkı sağlamayı amaçlamaktadır.

Bu çalışmada KÇ yöntemiyle yapılan mikro-analizler, ilk analizlerin problemli sınav maddelerinin ele alındığı etkileşimlere yönelik belirlediği üç aşamayla ilgili ayrıntılı bulgular ortaya çıkarmıştır. Problemi belirleme (PB), değişiklik önerme (DÖ), ve değişikliği gözden geçirme (DG) aşamaları bir dil sınavı maddesinin öğretmenler tarafından olumsuz olarak değerlendirildiği Mİ etkileşimlerinin genel yapısal düzenini oluşturmaktadır. Dolayısıyla, bu durumlarda incelemeye sunulan sınav maddeleri problemleştirilmekte, düzeltilmekte ve değiştirildikten sonra tekrar incelenmektedir. İlk aşama (PB) katılımcıların maddede tespit ettikleri problemi duyurmaları ve bu konuyla ilgili olan konuşmalarını kapsamaktadır. Mİ grubu ortaya atılan konuyla ilgili görüş birliğine vardığı zaman bu aşama sona erer ve eğer grup problemi kabul ederse ikinci aşama olan DÖ başlar çünkü artık bunu takip eden bir eylem gerekli kılınmıştır. Bu ikinci aşama ilk aşamada problemleştirilen sınav maddesini düzeltmek ve gereken değişiklikle ilgili ortak karara varmak için öneride bulunma dizilerinden oluşmaktadır. Grup bir değişikliğe karar verdiği anda sınavcı değişikliği sınav kâğıdının üzerine not düşer. Katılımcılar bu sefer düzeltilmiş maddeye uygunluğunu bir kez daha kontrol etmek için yönelirse, son aşama olan DG gerçekleşir ve grup değişikliği onaylayıp maddeye son halini verdiği Mİ etkileşimi nihayetinde sona ermiş olur.

Bulgular Mİ etkileşimlerinde tespit edilen üç aşamanın yerel olarak yönetildiğini ve farklılıklara tabi olduklarını göstermiştir (Drew ve Heritage, 1992). Bir takım şartlara bağlı olarak bazı aşamalar atılabilmekte, daha önce tamamlanmış olan aşamalar yeniden gerçekleşebilmekte ya da aynı aşamanın içindeki etkileşimsel desenler hedefe ulaşılan kadar tekrarlayabilmektedir. Ayrıca aşamalar arasında dönüşlü bir bağıntı (reflexive relation) olduğu görülmüştür (Robinson, 2013). Bir aşamanın oluşması bir önceki aşamanın gerçekleşmesine ve bunun oluşturduğu şartlara bağlı olarak ortaya çıkmaktadır. Örneğin, DÖ aşaması sadece katılımcılar problemi kabul ettikleri takdirde zorunlu kılınmaktadır.



Aynı şekilde DG aşaması bir önceki aşamada maddede herhangi bir değişikliğin yapılması koşuluyla gerçekleşebilmektedir. Dolayısıyla, Mİ etkileşimlerinin genel yapısal düzeni içinde her aşama belli bir amacı gütmekte, Mİ görevini etkileşim içinde ve etkileşim aracılığıyla başarılı bir şekilde tamamlama sürecinde katılımcılar tarafından ortaklaşa oluşturulmaktadır (Heritage, 2005; Heritage ve Clayman, 2010). Genel yapısal düzen, her bir aşamasının tamamlanmasıyla aktivitenin ilerlemesini ve birbiriyle ilişkili olan eylemlerin yansıtılabilirliğini sağlamaktadır (Robinson, 2013). Bu anlamda analizlerin ortaya koyduğu genel yapısal düzen, eğitim ortamlarında öğretmenler arasında gerçekleşen dil sınavı Mİ etkileşimlerinin kurumsal yapısını göstermektedir (Drew ve Heritage, 1992). Mİ etkileşimlerinde hedef, sınav maddelerinin içerik ve biçim bakımından belirlenen amaç için uygunluğunun incelenmesi olduğu için öğretmenler maddelerin bu açıdan bir değerlendirmesini sunmakta, daha sonra kurumsal sorumlulukların bir parçası olarak bunun da ötesine geçip problemleştirilmiş maddeleri herkesin kabul edeceği şekilde düzeltmektedirler. Bu da maddede bir problem ortaya çıktığında katılımcıların hareketlerini konuşma içinde ve konuşma aracılığıyla nasıl şekillendirdiklerini ve Mİ oturumlarında kendilerinden beklenen kuruma özgü görevleri etkileşim içinde açıkça nasıl gerekli kıldıklarını göstermektedir (problemi açıklama, bir değişiklik önerme, değişikliğe karar verme, değişikliği gözden geçirme).

Yapılan mikro-analizler, ortaya çıkan üç aşamanın her birinin genel yapısal düzeni için modeller sağlamış, sınav maddesine yönelik karşılıklı anlaşılabilirliği ve ortak karar vermeyi sağlayan yaygın etkileşimsel eylemleri ortaya çıkarmıştır. Dil sınavı Mİ etkileşimlerinde her bir aşamanın hedef yöneliminden dolayı, bu aşamaları tanımlayan temel bir eylem söz konusudur: PB aşamasında maddede tespit edilen problemle ilgili açıklamada bulunmak, DÖ aşamasında öneride bulunmak, DG aşamasında ise değiştirilmiş madde için grubun onayını istemek. Tespit edilen etkileşimsel eylemlerden bazıları her bir aşamaya özgü iken (örn., PB aşamasında problemle ilgili ön soruşturma gerçekleştirme; DÖ aşamasında bir değişiklik önerme, ortak metin üretme, kendisine katılmama, kararı sınav kâğıdına yazma; DG aşamasında değiştirilmiş madde için onay isteme), diğerleri Mİ görevi boyunca her bir aşamanın amacına hizmet edecek şekilde tekrarlanmaktadır.

Bu eylemlerden açıklama yapma, sorgulama, çeşitli etkileşimsel kaynaklar kullanarak (örn., tahmini tamamlama, değiştirilmiş tekrar, reformülasyon) karşılıklı anlaşılabilirlik ve görüş birliği gösterme tüm aşamalarda görülmekte, incelenen sınav maddesi için önemli dizisel

sonuçlar ve uygulamalar ortaya koymaktadır. Bu eylemleri yürütürken katılımcılar öğretmenlik tecrübeleri (örn., geçmiş öğretim olayları, öğrencilerin ikinci dil bilgi ve becerileri), sınav hazırlama ve dil uzmanlıklarından yararlanarak ve mevcut ortak kaynaklardan kanıt sunarak etkileşim içindeki konuşmaya profesyonel yetilerini getirmektedirler. Bu bakımdan, bu çalışma bu tür gerekçelere dayanarak yapılan açıklamaların (accounts) sadece problemi, öneriyi ya da maddede yapılan değişikliği desteklemek için değil, aynı zamanda bir iddiayı sorgulamak, görüş birliği sergilemek ya da grupta ortak anlayışı sağlamak için yapıldığını göstermiştir. Bu da dil sınavı Mİ etkileşimleri gibi işbirlikçi işyeri etkileşimlerinde açıklamaların proaktif, yaratıcı ve sorun çözme işlevini göstermektedir (Firth, 1995).

Söz konusu eylemler (açıklama yapma, sorgulama, karşılıklı anlaşılabilirlik ve görüş birliği gösterme) problemin, önerilen değişikliğin ve değiştirilmiş maddenin çok yönlü açıdan işbirliği içinde tartışılmasını ve görüşülmesini sağlamakta, bu ise problem-odaklı bir bağlamda ortaya atılan konuların daha iyi anlaşılmasına neden olmaktadır. Diğer bir ifadeyle, katılımcılar argümanlarını ortaklaşa oluşturmakta (Smithson ve Diaz, 1996) ve söz konusu Mİ etkileşimini sonlandırmadan önce sınav maddesi, önerilen değişiklik ya da düzeltilmiş madde üzerinde dikkatli düşünmektedirler. Bununla ilgili olarak, katılımcılar bir takım etkileşimsel kaynaklar kullanarak tartışılan konulara yönelik (tespit edilmiş problem, önerilen alternatifler, maddede yapılan değişiklikler) işbirlikçi sorgulamanın önünü açmaktadırlar. Birçok kesitte görüldüğü gibi PB ve DÖ aşamalarında fikirleri kesinlik ifadeleriyle değil de derecesi düşürülerek (downgrade) öne sürme (Koester, 2002), DÖ ve DG aşamalarında şeytanın avukatlığını yapma (Davidson ve Lynch, 2002; Leyland vd., 2016; MacDougall ve Baum, 1997; Percy vd., 2015), PB aşamasında bir ön-genişletme dizisinde problemle ilgili ön soruşturma başlatma (Pomerantz ve Heritage, 2013) ve bütün aşamalarda görüldüğü gibi ele alınan konuları sorgulama katılımcıların etkileşim içinde işbirlikçi sorgulamayı sağlamada kullandıkları eylemlerdir. Ayrıca, yaygın bir yaklaşım olarak, grupta bir anlaşma sergilenmesine rağmen katılımcılar problemi, önerilen değişikliği ya da değiştirilmiş maddeyi minimal olmayan bir art-genişletme dizisinde (Schegloff, 2007; Stivers, 2013) daha ayrıntılı olarak tartışmaktadırlar.

PB ve DÖ aşamalarında birbirlerinin iddialarını sorgulamanın ve bunlara karşı çıkmanın sonucunda ortaya çıkan anlaşmazlığın yönetimine bakıldığında, katılımcıların görüş değişikliği ya da kendi görüşünden vazgeçme yoluyla anlaşmaya doğru bir yönelim

sergiledikleri görülmüştür (Sharma, 2012; Wasson, 2000). Katılımcıların PB aşamasında problemi sorguladıktan ve daha sonra diğer katılımcılar tarafından karşı sorgulamaya tabi tutulduktan sonra problemi kabul etmeleri ya da DÖ aşamasında diğer katılımcılar tarafından açık bir kanıtla sorgulandıktan sonra teklif edilen bir önerinin kabul edilmemesi konusunda anlaşmaları bunun en belirgin örnekleridir. Katılımcılar anlaşmazlıklardan çıkmak için belli ön-kabul ediş eylemleri (örn., önceki açıklamayı yeniden tasarlama, karşı sorgulamada bulunarak daha fazla açıklama yapma ve kanıt sunma) gerçekleştirmekte ve böylelikle mevcut görev ile ilgili sonraki eylemin yerine getirilebilmesine fırsat tanımaktadırlar (Sharma, 2012). Ele alınan konuların açıklama ve sorgulama yoluyla tartışılıp değerlendirilmesi sonucunda her bir aşamada sınav maddeleriyle ilgili ortaklaşa kararlar alınmaktadır. PB aşamasında Mİ grubu incelenen maddenin düzeltilmesini, atılmasına veya olduğu gibi tutulmasına karar vermekte, ya da bu aşama duyurulan problemin yeniden tanımlanmasıyla son bulmaktadır. DÖ aşamasında katılımcılar önerilen değişikliği reddetmekte veya kabul etmekte, ya da öneriyi sunduktan hemen sonra kendi önerilerini geri çekmektedirler. DG aşamasında ise grup değişikliği ya onaylamakta ya da bu değişikliğe karşı sunulan itirazı kabul ederek değiştirilmiş maddede ortaya çıkan yeni problemi çözmek için DÖ aşamasını yeniden gerçekleştirmektedirler. Bu da maddede yapılan ana değişikliğin başka değişikliklere neden olabileceğinden dil sınavı Mİ etkileşimlerinin dikkatli ve eleştirel bir bakışla devamlı inceleme ve düzeltme gerektirdiğini göstermektedir.

PB ve DÖ aşamalarında anlaşmazlık eylemi için farklı yeğleme (preference) durumları sergilenmiştir. PB aşamasında anlaşmazlık genelde yeğlenmeyen bir yanıt olarak ele alınırken, DÖ aşamasında anlaşmazlığın ekstra bir etkileşimsel çaba gerektirmeksizin kabul edilir ve belirtilmez (unmarked) olduğu bulunmuştur. Bu durum, aşamaların hedef yönelimleri ve değerlendirme eyleminin Mİ etkileşimlerinin ayrılmaz bir parçası olması ile açıklanabilir (Koester, 2006; Lindström ve Mondada, 2009). Amacının maddeyi dikkatli bir şekilde değerlendirip olası problemleri tespit etmek olduğu PB aşamasında, itirazlar incelenen sınav maddesinin olumsuz değerlendirilmesine yönelik sunulmaktadır. Diğer tarafta, amacının problemleştirilmiş sınav maddesini düzeltmede olası en iyi çözümü aramak olduğu DÖ aşamasında, itirazlar bir öneriye yönelik yapılan olumsuz değerlendirmelerin kendisini oluşturmaktadır. Bu aşamalarda anlaşmazlığın yönetilme şekliyle anlaşılacağı gibi görevin yerine getirilmesi için katılımcıların ya sıralı çiftin ilk kısmı (PB aşaması) ya da sıralı çiftin ikinci kısmında (DÖ aşamasında öneriye yanıt kısmı) açıklama ve kanıtlarla sunulan olumsuz değerlendirmeleri kabul etmeleri beklenmektedir. Sonuç olarak, PB aşamasında

katılımcıların sınav maddesiyle ilgili olumsuz değerlendirmelerine itiraz edildiğinde ki bu da genellikle bir hafifletme (mitigation) yöntemiyle ifade edilmektedir (örn., anlaşmazlığa anlaşma ifadeleriyle giriş yapmak), grup üyeleri sınav maddesinde tespit edilen problemin genellikle kabul edilmesine yönelik etkileşimsel bir çaba göstermektedirler (örn., dizi genişlemeleri). Böylelikle, sunulan probleme (maddenin olumsuz değerlendirilmesine) katılmama eylemine yeğlenmeyen bir yanıt olarak yönelmektedirler. Öte yandan, DÖ aşamasında teklif edilen bir öneriye itiraz edildiğinde, grup üyeleri genellikle önerileriyle ilgili olumsuz değerlendirmelere karşı çıkmayıp olası en iyi çözüm arayışlarına devam etmekte ve böylelikle önerilerine verilen katılmama yanıtlarına yeğlenmeyen bir eylem olarak yönelmemektedirler.

Bununla birlikte, PB ve DÖ aşamalarında anlaşmazlık için farklı yeğleme durumları gösterilse de bulgular ayrıca sorgulama eyleminin ve karşı duruş sergilemenin özellikle katılımcıların anlaşmazlıklara katılıp bir görüş ya da öneriye karşı sunulan itirazları kabul ettikleri durumlarda hareket biçimini farklı şekillerde değiştirebildiğini göstermiştir. Diğer bir ifadeyle, Mİ grubunda sorgulama yoluyla sergilenen anlaşmazlık sınav maddesine yönelik alınan kararlarda ve grupta karşılıklı anlaşılabilirliği oluşturmada önemli bir rol oynamaktadır. Bu da dil sınavı Mİ etkileşimlerinde bir konunun farklı bakış açılarıyla tartışılmasının anlaşma ve ortak anlayış çerçevesinde yönetildiği sürece ne kadar önemli olduğunu göstermektedir (Myers, 1998). Dolayısıyla, anlaşmazlığın kurumsal ortamlarda iyi değerlendirilmemiş kararları önleyen görev destekleyici bir eylem olduğu bulunmuştur (Angouri, 2012; Kangasharju, 2002).

Bu çalışmanın bulguları, meslektaşlar arasında gerçekleşen Mİ uygulamasında katılımcıların sadece etkileşimsel hedeflere değil aynı zamanda ilişkisel hedeflere yöneldiklerini ortaya koymuştur (Koester, 2006). Örneğin, katılımcılar bir maddedeki problemi açıklarken, probleme karşı çıkarken ya da problemleri maddeyi düzeltmek için herkes bir alternatif arayışı içerisindeyken bir öneri teklifinde bulunurken belirsizlik sözcükleri (hedges), açıklama, gecikme araçları, anlaşmazlıklardan önce kullanılan anlaşma ifadeleri gibi hafifletme (mitigating) kaynakları kullanmaktadırlar. Böylelikle, katılımcılar etkileşim içindeki diğer kişilere karşı nezaket sergilemekte, yüz durumları (Brown & Levinson, 1978; Goffman, 1955) ve diğerleriyle olan ilişkilerini (Heritage, 1988) dikkate almaktadırlar.

Bu çalışmada öğretmenlerin sınav maddelerini işbirliği içinde inceleyip düzeltirken konuşma içinde ve konuşma aracılığıyla çeşitli sınav hazırlama ilkelerine yönelim gösterdikleri bulunmuştur. Öğretmenler literatürde genel olarak madde yazımı için, daha belirgin olarak da seçilmiş yanıt (örn., çoktan seçmeli, eşleştirme) ya da sınırlı üretim (örn., boşluk doldurma) madde biçimleri için tanımlanmış olan sınav hazırlama ilkelerinin (örn., Carr, 2011; Fulcher, 2010; Haladyna ve Rodriguez, 2013; Harris, 1969; Hatipoğlu, 2009; Heaton, 1990) birçoğuna yönelmektedirler. Katılımcılar bu ilkelere incelenen maddeyle ilgili bir konuyu açıklama ve sorgulama esnasında başvurumaktadırlar. Bu durumlarda, yönelinen ilkelerin sınav maddesi tarafından ya ihlal edildiği ya da yerine getirildiği iddia edilmekte, bu ilkelere uyulmama durumunda ise katılımcılar problemi çözme yönünde bir adım atmaktadırlar. Mİ etkileşimlerinde gönderme yapılan sınav hazırlama ilkeleri şu konuları kapsamaktadır; (1) öğretme ve öğrenme bağlamı dikkate alındığında madde içeriğinin önemi ve geçerliliği, (2) maddenin cevabı (örn., “tek doğru cevap” ilkesi), (3) öğrencilerin ikinci dil bilgi ve becerileri (örn., “zorluk derecesi” ilkesi), (4) maddede kullanılan dil, (5) seçeneklerin uygunluğu, okuma parçasında kullanılan kelimelerin basitliği, bağlamın yeterliliği, maddenin bağımsızlığı gibi sınav hazırlamayla ilgili belli konular. Bu da katılımcıların sadece öğretmenlik tecrübelerine değil, aynı zamanda sınav hazırlama ve dil bilgilerine dayanarak dil sınavı maddelerini çeşitli açılardan analiz edip değerlendirdiklerini göstermektedir. Böylelikle, katılımcılar daha büyük ilkelere gönderme yaparak iddialarını makul ve profesyonel bir temele dayandırmaktadırlar (Waring, 2017).

Dil sınavı Mİ etkileşimlerinin tüm aşamalarında katılımcıların kurumsal rol dağılımına bir yönelim sergilenmiştir. PB aşamasında konuşma içinde ve konuşma aracılığıyla gerekli kılınan rollerin daha belirgin olduğu görülmüştür. Çünkü incelenen ürün sınavcıya aittir ve burası öğretmenlerin Mİ etkileşimlerinde esas amaçları olan sınav maddesini değerlendirme hakkına sahip oldukları aşamadır (Heritage, 2013). Bu sebeple, PB aşamasında Mİ görevini gerçekleştiren öğretmenler, açıklama yaparak ve kanıt sunarak sınav maddesini problemleştirmekte, sınavcılar ise beklenen doğru cevabı bildirme, yansıtılan problemin anlaşılmasını sağlamak için maddenin cevabına yönelik önden ayrıntılı açıklamada bulunma, ürünlerindeki problemde sorumlu olduklarından problemi sorgulama gibi eylemleri yerine getirmektedirler. Bununla birlikte, bu rol dağılımı daha sonraki aşamalarda kısmen ortadan kalkmaktadır. Çünkü bu aşamalarda grup maddeyi işbirlikçi bir şekilde ortak çalışmanın ürünü olarak yeniden üretmektedir. Bu aşamalarda katılımcıların görev ve sorumluluklarının PB aşamasından farklı olarak sabit olmadığı ve birden fazla aşamada gerekli kılındığı

görülmüştür. Örneğin, PB ve DÖ aşamalarında, madde incelemeyi yapan öğretmenler sınava girme sürecine (Alderson vd., 1995), sınavcılar ise madde yazımı sürecine gönderme yapmaktadırlar. Ayrıca, üç aşamanın hepsinde madde incelemeyi yapan katılımcılar iddialarını öğrenci ve öğretim açılarından ifade edip, öğretim ve sınav arasındaki bağlantıyı oluşturarak öğretmen olarak kurumsal rollerine yönelmektedirler. Sınavcılar ise PB aşamasında açıklanan probleme, DÖ ve DG aşamalarında da kabul edilen değişikliklere yönelik doğru ve ortak anlayış sağlama sorumluluklarına yönelmektedirler.

Katılımcıların kurumsal rol dağılımına bir yönelim sergilenmesine rağmen, analizler Mİ görevinde katılımcıların ortak sorumluluk ve eşit katılımlarını gösteren açık kanıtlar da ortaya koymuştur. Katılımcılar böyle bir görev-odaklı işbirlikçi işyeri etkileşiminde ortaklık kurmaktadırlar. Bunun en belirgin kanıtları katılımcıların işbirlikçi etkileşimsel eylemleri, kapsayıcı dil kullanımı (örn., kapsayıcı “*biz*”) (Sacks, 1992), bakış hareketleri ve sözsüz eylemleridir (örn., metni gözden geçirme, kalemle kelimelere işaret etme, düşünen yüz ifadesi, bakışı metin ve madde arasında kaydırma). Katılımcılar konuşma, tekrar eden etkileşimsel yapılar, karşılıklı erişimleri bulunan yazılı materyaller, bakış, yazma ve okuma gibi çoklu kaynakları koordine ederek karşılıklı yönelim oluşturmakta, değerlendirilmekte olan sınav maddeleri üzerinde ortak ilgi odağı elde etmektedirler (Balaman ve Sert, 2017; Ford, 1999; Goodwin, 2000, 2007; Lindström ve Mondada, 2009). Mİ sürecine tüm katılımcıların aktif bir şekilde dahil olması tartışılan konulara eşit epistemik erişimlerinin olması ve epistemik bir topluluk oluşturmalarından kaynaklanmaktadır (Heritage, 2013).

Bu çalışmanın bulguları, literatürde tespit edilmiş boşluklara katkı sağlayarak dil sınavı oluşturma alanına, daha belirgin olarak Mİ aşamasına yönelik uygulamalar sunmaktadır. İlk olarak, dil sınavı oluşturma modellerinde Mİ önemli adımlardan biri olarak ele alınmış ve bu aşamada öğretmenler gibi ilgili kişilerle gerçekleştirilen etkileşimlerin önemi vurgulanmıştır (örn., Bachman ve Palmer, 1996; Carr, 2011; Fulcher, 2010). Ayrıca, belirgin olarak Mİ uygulamasına odaklanmamalarına rağmen, dil sınavı oluşturmada öğretmen geribildirimi ve ilgili kişilerle olan etkileşimin rolü üzerine birkaç çalışma bulunmaktadır (örn., Davidson ve Lynch, 2002; Green ve Hawkey, 2011; Kim vd., 2010; So, 2014; Tan ve Turner, 2015), fakat araştırmacının bilgisi çerçevesinde, dil sınavı oluşturmada etkileşim ve Mİ uygulamasına verilen öneme rağmen dil sınavı hazırlama ve Mİ sürecinde öğretmenlerle gerçekleşen etkileşimlerin mikro-ayrıntıları çalışılmamıştır. Bu çalışma şunları yaparak bu boşluğa yönelmektedir; (1) doğal gelişen veriyle çalışıp, içeriden bir bakış açısıyla Mİ

etkileşimlerinin mikro-ayrıntılarını sistemli bir şekilde inceleme olanağı sağlayan KÇ yöntemini kullanma, (2) müfredata dayalı dil sınavı Mİ etkileşimlerinin aşamaları için bir model geliştirip, bununla birlikte tekrar eden etkileşimsel eylemlerin, konuşma içinde ve konuşma aracılığıyla yönelim gösterilen sınav hazırlama ilkeleri ve katılımcıların kurumsal rol dağılımının bir açıklamasını sunma. KÇ yoluyla yapılan böylesi ayrıntılı mikro-analitik bir açıklama, sınav oluşturma ve etkileşim arasındaki yakın ilişkiyi göstermektedir.

İkinci olarak, bu çalışma büyük ölçekli sınavlara kıyasla ihmal edilmiş olan yerel ortamlardaki müfredata-dayalı dil sınavlarının etkileşim içinde oluşturulmasına dikkatleri çekmekte, böylelikle de öğretmenlerin günlük öğrenci değerlendirme uygulamalarını ortaya çıkarmaktadır. Bu çalışma güncel dil sınavı çalışmalarındaki süreç-odaklı nitel çalışmalara, daha belirgin olarak da çoğunlukla konuşma sınavlarındaki dil yeterlik etkileşimleriyle ilgilenen KÇ çalışmalarına katkı sağlamaktadır (örn., Galaczi, 2014; Seedhouse, 2013; Young ve He, 1998). Bu açıdan bakıldığında, bu çalışma oldukça soyut kurumsal kavramların ötesine geçip öğretmenlerin günlük işlerinin bir parçası olarak yerel ortamlarda gerçekleştiği şekliyle dil sınavı hazırlamaya yönelik süreç-odaklı çalışmalara olan ihtiyacı karşılamaktadır. Üçüncü olarak, bu çalışma Mİ uygulamasının literatürde önerildiği gibi önceden belirlenmiş adım adım standart kontrol listeleri ve alt işlemleri (örn., Bachman ve Palmer, 1996; Brown, 2001; Fulcher ve Davidson, 2007) takip ederek basit bir şekilde gerçekleştirilemeyeceğini göstermiştir. Çünkü Mİ etkileşiminin aşamaları arasındaki dönüşlü bağıntı ve incelenen sınav maddelerine yönelik ortaya çıkan farklı etkileşimsel dizi ve sonuçlardan anlaşılacağı üzere Mİ görevinin uygulamada özyineli (iterative) ve dinamik bir süreç içerdiği bulunmuştur.

Bu çalışmanın bulguları, dil sınavı oluşturmaya yönelik metin ve çalışmalardaki sabit görüşleri desteklemekle birlikte Mİ uygulamasının öğretmenlerin işyeri etkileşimleri içinde belirgin bir şekilde nasıl gerçekleştiği konusunda yeni bakış açıları sunarak dil sınavı oluşturmada Mİ uygulamasının önemini göstermektedir. İlk olarak, bu çalışma, ne kadar tecrübeli ve yetenekli olsa da maddeyi yazan kişinin sınav maddelerinde gözünden kaçan çeşitli problemlerin olabileceği ve meslektaşlarla etkileşim içinde gerçekleşen Mİ uygulamasının kusurlu maddeleri işbirlikçi bir şekilde tartışma yoluyla tespit edip, bunlara çözüm getirme konusunda etkili bir yöntem olduğu görüşlerini desteklemektedir (örn., Alderson vd., 1995; Bachman ve Palmer, 1996; Baranowski, 2006; Brown, 1996; Downing, 2006; Fulcher 2010; Fulcher ve Davidson, 2007; Hughes, 1989; Linn, 2006). İkinci olarak,

bu çalışma meslektaşlardan etkileşim içinde ve etkileşim aracılığıyla geribildirim almanın geçerlik argümanı (validity argument) için önemli bir ön kanıt sağladığı bulgusuna yönelik gerçek Mİ etkileşimlerinden kanıt sunmaktadır (Green ve Andrade, 2010; Lynch ve Davidson, 1994; So, 2014; Tan ve Turner, 2015). Böylelikle, bu çalışma geçerlik kanıtının (validity evidence) Mİ uygulaması yoluyla ilgili kişilerden etkileşim içinde ve etkileşim aracılığıyla nasıl elde edilebileceğini canlı biçimde göstermekte, madde geçerlik saptama işleminin nitel olarak nasıl gerçekleştiğini açıklamaktadır (Downing, 2006; Weir, 2005; Wolf vd., 2008).

Üçüncü olarak, bu çalışma öğretmenlerin Mİ uygulamasına dahil olmalarının ve etkileşimlerinin önemini göstermektedir (Alderson vd., 1995; Brindley, 2001; Dikilitaş vd., 2018; Fulcher ve Davidson, 2007; Xerri ve Vella Briffa, 2018). Öğretmenlerin geçmiş öğretim olayları ve öğrencilerin ikinci dil bilgi ve becerilerine gönderme yaparak öğretim ve sınav arasındaki uyumsuzlukları tespit etmede (Lynch ve Davidson, 1994), böylelikle de yerel ihtiyaçları karşılayan “ekolojik olarak hassas” bir sınavın geliştirilmesine yardımcı olabilecekleri görülmüştür (Fulcher, 2010). Ayrıca, öğretmenler sınav maddelerini sınav hazırlamayla ilgili çeşitli konuları da göz önünde bulundurarak incelemekte, madde geçerliği saptama sürecine önemli katkılar sağlamaktadırlar. Sonuç olarak, bu çalışma, öğretmenlerle olan Mİ etkileşimlerinin dizisel yapısının ve belli etkileşimsel eylemlerin kullanımının sınav maddelerini nasıl şekillendirdiğini, karar verme sürecini nasıl etkilediğini ve bu maddelerin geçerliğinin saptanmasına nasıl katkı sağladığını göstererek dil sınavı oluşturmada Mİ aşamasının sosyal bir uygulama olarak daha iyi anlaşılmasını sağlamakta ve buna yönelik yeni bakış açıları kazandırmaktadır.

Bu çalışmanın bulguları, Mİ görevinin tamamlanması için gerekli olan etkileşimsel eylemleri, katılımcıların sınav hazırlama ilkeleri ve kurumsal rol dağılımlarına yönelimleriyle birlikte açıklayarak öğretmenlerin mesleki yeterliliklerine yönelik uygulamalar sunmaktadır. Kurumsal ortamlarda daha geniş sosyal becerilerin temel unsuru olarak etkileşimsel yeti, insanların eğitim ve profesyonel hedeflerini gerçekleştirmelerini mümkün kılmakta ve çeşitli kurumsal dünyalara erişimlerini sağlamaktadır (Pekarek Doehler ve Petitjean, 2017). Bir eğitim ortamında öğretmenler tarafından gerçekleştirilen görev-odaklı dil sınavı Mİ uygulamasındaki etkileşimsel yetileri KÇ yoluyla ortaya çıkararak, bu çalışmanın iş dünyasına yeni adım atmış öğretmenlere yeni etkileşimsel ortamlardaki taleplerin üstesinden gelme konusunda yardımcı olması beklenmektedir



(Pekarek Doehler ve Petitjean, 2017). Bunun yanı sıra, bağlama özel bir araştırma ortamında belli bir kurumsal görev olan dil sınavı Mİ eylemine odaklanmasına rağmen bu çalışmanın belli görevlerin meslektaş etkileşimi içinde ortaklaşa yerine getirilmesi gerektiği profesyonel bağlamlar için uygulamalar sunduğu söylenebilir.

Bu nedenle, bu çalışma aynı zamanda eğitim ortamlarında öğretmenlerin görev ve hedef odaklı işyeri etkileşimlerinin anlaşılmasına katkı sağlamaktadır (örn., Barnes, 2007; Van Kruiningen, 2013; Verkuyten, 2000). Araştırmacının bilgisi çerçevesinde kurumsal bağlam olarak bir eğitim ortamında dil sınavı oluşturmayla ilgili öğretmenler arasında geçen etkileşimler üzerine KÇ çalışması bulunmamaktadır. Fakat okullarda öğretmenlerin dahil olduğu farklı işbirlikçi işyeri etkileşimlerinin mikro-analitik ayrıntılarını inceleyen birkaç çalışma yapılmıştır. Bu çalışmalar, öğrencilerin etkileşimsel yetilerini değerlendirme (Sandlund & Sundqvist, 2019), bir üniversite dersinin yeniden geliştirilmesi (Van Kruiningen, 2013), müfredat geliştirme (Barnes, 2007), başarısız okul karne notlarının tartışılması (Verkuyten, 2000), öğrenci davranışını yönetmek için kurulmuş bir sistemin kurallarını gözden geçirme (Baker, 1997), bir okul-üniversite ortaklığının oluşturulması (Hall ve Danby, 2003) gibi farklı kurumsal hedef ve görevlere yönelik öğretmen etkileşimlerini kapsamaktadır. Bu çalışmalarda KÇ kullanılarak işyeri temelli öğretmen buluşmalarında birtakım etkileşimsel desenler ortaya çıkarılmıştır. Eğitim kurumlarındaki bu çalışmalarda, örneğin, söz sırası tamamlama (Van Kruiningen, 2013), formülasyon (Barnes, 2007; Van Kruiningen, 2013), açıklama (Baker, 1997; Verkuyten, 2000), “uzman olma” gibi üyelik uyamlama (membership categorization) araçlarının kullanımı ve oluşturulması (Baker, 1997; Hall & Danby, 2003) gibi etkileşimsel kaynakların analizleri sunulmuştur. Bu çalışmaların görev tamamlama açısından farklı bir amacı olsa da bunların ve mevcut çalışmanın KÇ bulguları arasında kurumsal ortamlarda gerekli olan etkileşimsel yeti açısından benzerlikler bulunmaktadır. Bu da genel anlamda eğitim ortamlarında öğretmenlerin meslektaşlarıyla gerçekleştirdikleri işbirlikçi görev-odaklı etkileşimlerle ilgili farkındalık oluşturmaktadır. Dolayısıyla, bu çalışmanın aynı zamanda öğretmenler arasında gerçekleşen kurumsal konuşmalara yönelik yapılmış KÇ çalışmalarına katkı sağlaması beklenmektedir.

Son olarak, bu çalışmanın bulguları ikinci dil öğretmen yetiştirme programları için uygulamalar sunmaktadır. Sınav maddelerinin olumsuz değerlendirildiği dil sınavı Mİ etkileşimleri için veriye dayalı bir model (bkz. Şekil 7) geliştirerek bu çalışmanın öğretmen

ve öğretmen adaylarına kurumlarında böyle bir görevin gereksinimlerini karşılama konusunda yardımcı olması beklenmektedir. Bu çalışmada önerilen dil sınavı Mİ yetisi, İngilizceyi yabancı dil olarak öğreten öğretmenlerin değerlendirme okuryazarlıklarının ve kurumsal sorumluluklarının önemli bir parçasını oluşturmaktadır. Öğretmenlerin profesyonel bağlamlarında sınav hazırlama sorumluluklarını başarılı bir şekilde yerine getirebilmeleri için meslektaşlarıyla etkileşim içinde ve etkileşim aracılığıyla işbirliği içinde dil sınavlarını değerlendirme ve düzeltme becerilerini geliştirmeleri gerekmektedir. Dil eğitimi bağlamlarında Mİ eyleminin öğretmenlerin günlük sınav geliştirme uygulamalarının önemli bir parçası olduğu açık iken hizmet öncesi öğretmen yetiştirme programlarında öğretmen adaylarının etkileşim aracılığıyla arkadaşlarıyla birlikte sınav hazırladıkları, eleştirel bir şekilde bu sınavları değerlendirdikleri ve belli dil eğitimi ortamlarındaki gerçek sınav uygulamalarının öğretmen adaylarıyla paylaşıldığı uygulamalara ihtiyaç olduğunu gösteren çalışmalar bulunmaktadır (örn., Hatipoğlu, 2010; Kleinsasser, 2005; O'Loughlin, 2006; Stiggins, 2001). Yerel bir kurumsal bağlamda gelecek öğretmenlerin dahil olma olasılığı yüksek olan dil sınavı Mİ etkileşimlerinin mikro-analitik ayrıntılarını sunarak bu çalışmanın hizmet öncesi öğretmen yetiştirme programlarındaki bu ihtiyaçları karşılaması beklenmektedir.

Bu çalışmada ortaya çıkan etkileşimsel dil sınavı Mİ modeli, hem hizmet öncesi (İngilizce Öğretiminde Ölçme ve Değerlendirme Dersi) hem de hizmet içi öğretmen yetiştirme programlarında dil sınavı Mİ performansını takip etme ve değerlendirmede eğitim amaçlı bir rehber olarak kullanılabilir. Farkındalık kazandırmak için bu çalışmanın bulguları gerçek konuşma örnekleri ile birlikte ders materyali olarak kullanılıp öğretmen adayları ya da yeni başlayan öğretmenlerle paylaşılabilir. Bunu yaparken, bazı reflektif öğretmen eğitimi modelleri (örn., Sert, 2015) uyarlanıp, öğretmen adaylarının belirgin olarak dil sınavı Mİ görevi, genel olarak da işbirlikçi işyeri etkileşimleri için gerekli olan etkileşimsel yetilerini geliştirmede kullanılabilir. Böylelikle, öğretmen adayları dil sınavı uygulamalarında etkileşimin rolünü anlayıp, kurumsal bağlamlarda öğretmenlerin sınav bilgilerini ne zaman ve nasıl uyguladıklarını görebilirler ve ölçme-değerlendirme derslerinde öğrendiklerini anlamlılaştırabilirler. Ayrıca, bu çalışma Türkiye'de hizmet öncesi öğretmen yetiştirme programlarındaki İngilizce Öğretiminde Ölçme ve Değerlendirme ders sayılarının seçmeli ya da zorunlu ders olarak artırılması önermektedir. Böylelikle, bu derslerin müfredatında, dil sınavı Mİ uygulamasına yönelik KÇ bulgularının entegre edildiği, öğretmen adaylarının arkadaşlarıyla benzer Mİ etkileşimlerine dahil edildiği ve kendi etkileşimlerini

değerlendirmelerinin teşvik edildiği reflektif öğretmen eğitimi modellerinin kullanımı için yer ayrılabilir. Bu şekilde, ölçme-değerlendirme dersleri yerel ihtiyaçlara dönük daha uygun hale getirilebilir.

Bu çalışmada veri, İngilizcenin yabancı dil olarak öğretildiği belli bir yükseköğrenim kurumundaki sınav hazırlama bağlamından edinilmiştir. Fakat söz konusu bağlamı temsil etmesi açısından uzun bir zamana yayılan farklı öğretmen ve sınavcı grupları tarafından gerçekleştirilen dil sınavı Mİ oturumları kaydedilmiştir. Bu çalışmanın amacı genelleme yapmak olmasa da bulgular ve genel sonuçların Türkiye’de ve dünya genelinde sınavların yerel bir şekilde hazırlanıp sınıf öğretmenleri tarafından işbirlikçi bir şekilde incelenip birlikte oluşturulduğu diğer dil öğretimi programlarındaki sınav oluşturma uygulamalarına aktarılabilir olması umulmaktadır. Ayrıca, öğretmenlerin meslektaşlarıyla etkileşim içinde ve etkileşim aracılığıyla sınav maddelerini inceleyip düzelttikleri farklı eğitim ortamlarında da araştırmalar yapılarak bu çalışmanın bulguları genişletilebilir ve alana katkı sağlanabilir. Karşılaştırmalı bir yaklaşım teşvik edilerek bu çalışma daha ileriye götürülebilir. Örneğin, Mİ etkileşimleri (1) Türkiye’deki ve kültürler arası çeşitliliği ortaya koyabilecek diğer ülkelerdeki üniversitelerin İngilizce hazırlık okullarında, (2) ilkökul ve lise gibi eğitimin farklı düzeylerinde, (3) matematik, fen bilgisi gibi farklı derslerin müfredata-dayalı sınavları için incelenerek bulgular kıyaslanabilir. Bu çalışma KÇ kullanarak müfredata dayalı dil sınavı Mİ uygulaması için etkileşimsel bir model sunan ilk çalışma olduğu için etkileşim ve dil testi oluşturma arasındaki ilişkiyle ilgilenen diğer çalışmalar için de bir başlangıç noktası niteliğindedir. Benzer ya da farklı eğitim çevrelerinde Mİ uygulamasının etkileşimsel yapısıyla ilgili ampirik bulguların zaman içinde artmasıyla ortama özgü özellikler, benzer ve aykırı durumlar keşfedilebilir ve bu çalışmanın ortaya koyduğu Mİ görevinin temelini oluşturan yapı genişletilebilir (Arminen, 2005; Drew, 2003; Gardner, 2004; Heritage ve Clayman, 2010; Ten Have, 2007).

Bu çalışmada Mİ görevini yapan öğretmenlerin hepsi tecrübeli olduklarından, takip çalışmaları acemi öğretmenlerin meslektaşlarıyla dil sınavı Mİ etkileşimlerini nasıl yerine getirdiklerine de odaklanabilir. Bu öğretmenlerin etkileşim içinde ve etkileşim aracılığıyla Mİ performansları ve dil öğretiminde değerlendirme okuryazarlık gelişimleri boylamsal çalışmalar yoluyla izlenebilir. Ayrıca, eğitim ortamlarında öğretmenler arasında geçen görev-odaklı işyeri etkileşimlerinin farklı türleri de incelenebilir. Böylelikle, kurumlarındaki etkileşimsel gerekliliklerin üstesinden gelmek için öğretmenlerden beklenen etkileşimsel

yetiler bakımından bu alıřmanın bulgularının kıyaslanması mmkn olacaktır. Son olarak, bu alıřmanın bulgularının hizmet ncesi veya hizmet ii ğretmen yetiřtirme programlarında uygulandıėı K alıřmaları ve projeleri uygulamalı dilbilim arařtırma baėlamlarında gerekleřtirilebilir. Bunlar da İngilizce ğretiminde lme ve Deėerlendirme Dersleri ve profesyonel geliřim iin materyal hazırlamaya ynelik uygulamalar sunacak ve bu alıřmanın ğretmen eėitimine ynelik katkılarının daha iyi anlařılmasını saėlayacaktır.



## H. THESIS PERMISSION FORM / TEZ İZİN FORMU

### ENSTİTÜ / INSTITUTE

**Fen Bilimleri Enstitüsü/** Graduate School of Natural and Applied Sciences

☐

**Sosyal Bilimler Enstitüsü/** Graduate School of Social Sciences

☒

**Uygulamalı Matematik Enstitüsü/** Graduate School of Applied Mathematics

☐

**Enformatik Enstitüsü/** Graduate School of Informatics

☐

**Deniz Bilimleri Enstitüsü/** Graduate School of Marine Sciences

☐

### YAZARIN / AUTHOR

**Soyadı/** Surname : Can

**Adı /** Name : Hümeýra

**Bölümü/** Department : İngiliz Dili Öğretimi

### TEZİN ADI/ TITLE OF THE THESIS (İngilizce/ English) :

A Micro-analytic Investigation into EFL Teachers' Language Test Item Reviewing Interactions

**TEZİN TÜRÜ/ DEGREE:** Yüksek Lisans/ Master

☐

**Doktora/ PhD**

☒

1. **Tezin tamamı dünya çapında erişime açılacaktır./** Release the entire work immediately for access worldwide.

☐

2. **Tez iki yıl süreyle erişime kapalı olacaktır./** Secure the entire work for patent and/or proprietary purposes for a period of **two years**. \*

☐

3. **Tez altı ay süreyle erişime kapalı olacaktır./** Secure the entire work for period of **six months**. \*

☒

*\*Enstitü Yönetim Kurulu kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir.*

*A copy of the decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.*

**Yazarın imzası /** Signature..... **Tarih/** Date.....