TO THE MEMORY OF MY GRANDFATHER

TURKISH UNIVERSITY EFL STUDENTS' ORAL EXPRESSION OF CRITICAL THINKING IN CLASSROOM DISCUSSIONS

The Graduate School of Education

of

Bilkent University

by

Buket Esra Tarakçıoğlu

In Partial Fulfillment of the Requirements for the Degree of

MASTER OF ARTS

in

THE DEPARTMENT OF TEACHING ENGLISH AS A FOREIGN LANGUAGE

BILKENT UNIVERSITY

ANKARA

JULY 2008

THE GRADUATE SCHOOL OF EDUCATION

MA THESIS EXAMINATION RESULT FORM

July, 2008

The examining committee appointed by The Graduate School of Education

for the thesis examination of the MA TEFL student

Buket Esra Tarakçıoğlu

has read the thesis of the student.

The committee has decided that the thesis of the student is satisfactory.

Thesis Title:	Turkish University EFL Students' Oral Expression of Critical Thinking in Classroom Discussion
Thesis Supervisor:	Asst. Prof. Dr. JoDee Walters Bilkent University, MA TEFL Program
Committee Members:	Asst. Prof. Dr. Julie Mathews-Aydınlı Bilkent University, MA TEFL Program
	Asst. Prof. Dr. Valerie Kennedy Bilkent University, English Language and Literature

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Teaching English as a Second Language.

_____
(Asst. Prof. Dr. JoDee Walters)
Supervisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Teaching English as a Second Language.

_____
(Asst. Prof. Dr. Julie Mathews-Aydınlı)
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Teaching English as a Second Language.

_____
(Asst. Prof. Dr. Valerie Kennedy)
Examining Committee Member

Approval of the Graduate School of Education

_____
(Visiting Prof. Dr. Margaret Sands)
Director

ABSTRACT

TURKISH UNIVERSITY EFL STUDENTS' ORAL EXPRESSION OF CRITICAL THINKING IN CLASSROOM DISCUSSIONS

Buket, E. Tarakçıoğlu

M.A, Department of Teaching English as a Foreign Language

Supervisor: Asst. Prof. Dr. JoDee Walters

July, 2008

This study was designed to investigate the amount of the expression of critical thinking (CT) and the nature of L2 in tasks that require critical thinking in English for academic purposes (EAP) classrooms in Turkish universities.

This study was conducted with two advanced-level EAP classes, including 34 students in total, at Middle East Technical University, Turkey. The data was collected by video-taping the students' discussion in class and analyzing their transcripts in terms of critical thinking and nature of language.

The results indicated that students can express their critical thinking skills orally in L2 to a certain extent but they have difficulty in doing so as a result of their lack of necessary language. This study suggests the need for further work on oral expression of CT in EAP classrooms in Turkish universities in terms of the vocabulary and grammar students need to express themselves more accurately while engaged in tasks that invite CT in their oral L2.

Key Words: Critical thinking, EAP, spoken language, L2.

ÖZET

TÜRK UNİVERSİTELERİNDEKİ YABANCI DİL OLARAK İNGİLİZCE ÖĞRENCİLERİNİN SINIF İÇİ TARTIŞMALARINDA ELEŞTİREL DÜŞÜNCENİN SÖZEL OLARAK İFADESİ

Buket Esra Tarakçıoğlu

Yüksek Lisans, Yabancı Dil Olarak İngilizce Ögretimi Bölümü

Tez Yöneticisi: Yrd. Doç. Dr. JoDee Walters

Temmuz 2008

Bu çalışma yabancı dil olarak İngilizce öğretilen ortamlardaki akademik İngilizce derslerinde eleştirel düşüncenin ikinci dilde dışavurumunun miktarı ve eleştirel düşünceye davet eden sınıf içi çalışmalarda ikinci dilin yapısını araştırmak amacıyla tasarlanmıştır.

Bu çalışma Orta Doğu Teknik Üniversitesi'nde toplam 34 kişiden oluşan iki İngilizce 102 sınıfıyla gerçekleştirilmiştir. Gerekli bilgi bu öğrencilerin sınıfta yaptıkları bir tartışmayı videoya kaydedip, dökümlerinin eleştirel düşünce açısından incelenmesiyle elde edilmiştir.

Sonuçlar öğrencilerin eleştirel düşünme becerilerini bir dereceye kadar yabancı dillerine yansıtabildiklerini ama bunu yaparken gerekli dillerinin olmamasından dolayı zorlandıklarını göstermiştir. Bu çalışma akademik İngilizce sınıflarında sözel olarak eleştirel düşünceye davet eden çalışmalar sırasında öğrencilerin kendilerini daha doğru bir biçimde ifade edebilmeleri için ihtiyaçları olan kelime ve dilbilgisi açısından daha çok çalışılması gerektiğini önermektedir.

Anahtar Kelimeler: Eleştirel düşünce, Akademik İngilizce, konuşma dili, ikinci dil.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my thesis advisor, Asst. Prof. Dr. JoDee Walters for her invaluable guidance and patience throught the year. I would also like to thank to Julie Mathews-Aydınlı for her assistance and encouragement and Valerie Kennedy for revieving my thesis and providing me with invaluable feedback.

I owe my special thanks to the former director of Middle East Technical University, the Department of Modern Languages, Nihal Cihan, for giving me permission to attend the MA TEFL program and the current director Aylin Graves for her support with my studies. I would also like to thank my colleagues, Filiz Başaran and Meriç Gülcü and their students, who participated in this study at Middle East Technical Univerity.

Finally, I am grateful to my mother, granmother and especially my sister Banu for their continuous support and understanding through the year.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

Since the 1900s, approach to education has advanced from rote learning towards a more critical view of teaching and learning. The concept of critical thinking (CT) has been considered to be one of the overarching goals of education since the 1990s, especially in Western countries (Atkinson, 1997). It is not surprising that many English for Academic Purposes (EAP) courses at the university level frequently have CT underlying many of the objectives in their curricula as they aim to facilitate the adaptation of students to university. More specifically, EAP courses often aim to make critical thinking a cognitive habit for students in their receptive skills, namely reading and listening, and in productive skills, namely writing and speaking (Vermillion, 1997).

However, the relationship between CT and the speaking skill is an underrepresented area in the literature. Therefore, the purpose of this study is to investigate the amount of CT expression and the nature of students' second language (L2) while performing tasks that invite critical thinking. It will attempt to do this by analyzing students' language in class discussions requiring CT, according to a spoken CT framework and a set of criteria for examining the nature of language used that will be developed in the course of the study.

Background of the study

Discussions about critical thinking date further back than the 1990s. It has been with us for more than 2500 years, since the times of Socrates. Paul, Elder and Bartell (1997) state that Socrates introduced the idea of not believing in the value of ideas without asking profound questions to find clarity, logical consistency and adequate evidence first. This method is now known as 'Socratic Questioning' and is the most well-known and oldest critical thinking strategy. After Socrates established the tradition, many others examined the issue. From Plato, Aristotle, and the Greek skeptics, to John Dewey, Ludwig Wittgenstein and Piaget (20th century), many scholars have all contributed to critical thought and our knowledge of it in their own disciplines (Paul et al., 1997).

Today's understanding of critical thinking has been shaped by the aforementioned thinkers. However, we now have more detailed definitions of CT. Dewey, who is considered to be the 'father' of critical thinking, defines what he calls "reflective thinking", which is considered another term for CT now, as "active, persistent, and careful consideration of a belief or supposed form of knowledge in the light of the grounds which support it and the further conclusions to which it tends" (1910, p. 6). Another definition of critical thinking is given by Ennis (1989, as cited in Fisher, 2001, p. 4). He stated that it is "the process and skills involved in rationally deciding what to do or what to believe". In a more detailed attempt to define CT, Paul stated that the elements of CT are purposes, questions, points of view, information, inferences, concepts, implications and assumptions. He also stated that the standards to be applied to these elements are clarity, accuracy, relevance, logicalness, breadth, precision, significance, completeness, fairness, and depth

(2007). All these definitions signify a deep, rational and educated way of thinking, done to see beyond the apparent in all aspects of life.

The role of critical thinking in education is a topic of major discussion in the literature. Its place in education began to be widely discussed in the 1900s, with John Dewey's criticisms of the traditional education system, which was mostly based on rote learning, and his arguments for the need for a new, progressive system of education (1938). One of the most well-known educational philosophers who studied critical thinking as a part of the educational ideal is Bloom. His taxonomy of thinking skills, which involved knowledge, comprehension, application, analysis, synthesis, and evaluation (B. S. Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956), the latter three being the critical thinking skills, formed the basis of most subsequent studies. His model owes its success to the behavioral descriptions of each category. Although Bloom's taxonomy has been criticized by Ennis as being not really hierarchical (1981) and still vague (1993), it was clearer, more observable and more measurable than its predecessors.

Since the times of rote learning, much has been done in terms of what to change to adopt a critical approach to education. Collaborative learning is one of these methods. Collaborative learning entails:

> students at various performance levels work together in small groups toward a common goal. The students are responsible for one another's learning as well as their own. Thus, the success of one student helps other students to be successful. (Gokhale, 1995, p. 8)

Therefore, by nature, collaborative learning is a technique by which students' critical thinking skills are expected to be revealed as they are expected to collaborate and interact in achieving a goal. In her study, Gokhale (1995) concluded that both

collaborative and individual learning are equally effective in gaining factual knowledge. However, when the purpose of instruction is to enhance critical thinking, then collaborative learning helps achieve this aim better, through discussion, clarification of ideas, and evaluation of others' ideas. Newman, Webb and Cochrane (1995) also found evidence of enhanced critical thinking in L1 group work situations.

CT has also been studied extensively in terms of the acquisition of literacy skills. The importance of critical reading, as explained by Neilsen (1989), is that it causes people to gain independence of thought and action and this implies that people who lack critical reading skills will rely on others' thoughts. Writing is another skill which is directly related to critical thinking. Risinger (1987) points out that if designed appropriately, writing assignments are one of the most effective tools to improve CT. It is important to note that reading and writing skills are parallel to each other in terms of CT. Paul (1990) states that reading and writing are interrelated and parallel in that critical readers reconstruct the authors' ideas and critical writers construct their own ideas. Although Rasool, Banks and McCarthy (1993) and Barnet and Bedau (2002a) did not necessarily consider reading and writing as strictly interdependent, as Paul does, in their books reading and writing were studied together in the same broad CT framework.

The scholars who write about CT in spoken language mostly refer to it as language, which means they consider speaking and writing together. Fisher and Scriven (1997), for example, consider writing and speaking skills together as they are both productive skills, add presenting to these skills, and call this set of skills "critical communication" (p. 101).

To be able to say that CT is taking place in a certain situation, one should know what indicators to look for. Models of CT provide these indicators or specific behaviors, possibly by categorizing or describing them in stages. Theory-driven models to identify these indicators in L1 are very common in the literature. As previously mentioned, one of the most well-known CT taxonomies is Bloom's (B. S. Bloom, Thomas, & Madaus, 1971). Specifically for speaking and writing, Fisher and Scriven's (1997, pp. 103,104) model includes "veracity of claims and assumptions, soundness of inferences made or implied, suitability to audience, clarity of the presentation, comprehensiveness to the extent appropriate, conciseness to the extent appropriate, originality and power of the presentation". Research-driven models of CT indicators include both L1 and L2 studies. Kamin, Deterding and O'Sullivan (1998) investigated the amount and types of CT in third year medical students' face-to-face L1 discussions of cases. Their model is based on problem-solving, and consists of stages. Their stages are problem identification, problem description, problem exploration, applicability and integration (p. 221). Each stage has several sub-categories with samples of language. Another model has been developed by Newman, Webb and Cochrane (1995) to measure CT to compare online and face to face discussions. Their categories are relevance, importance, novelty, bringing outside knowledge/experience to bear on the problem, ambiguities, linking ideas/interpretation, justification, critical assessment, practical utility (grounding) and width of understanding (complete picture). A similar L1 model was developed by McLean (2005, pp. 10-11). He analyzed CT in student messages posted in computer conferences. The first dimension of his model is about the quality of CT and it includes clarity, relevance, depth, logic, preciseness, breadth, and support

elements. The second dimension is about the types of CT. The categories are clarification of the thesis, problem or question, making inferences and interpretations, supporting inferences and interpretations, and making value judgments. For L2, Uzuner (2007) designed a CT-indications model specifically for online discussions by ESL students. This model investigates CT in two different categories: educationally valuable talk and educationally less valuable talk. She exemplified each category with specific language uses. However, educationally valuable talk does not guarantee that there is CT. Yet, most of the categories are very much compatible with CT, such as exploratory, argumentational, critical, reflective, interpretive, analytical, informative and implicative talk.

Statement of the Problem

As the above discussion shows, the general categories of critical thinking in L1 have been studied widely in the literature (see for example J. Dewey, 1910; Paul, 1990). The general and linguistic categories have also been investigated extensively in reading and writing in L1 (see for example Barnet & Bedau, 2002b; Neilsen, 1989; Smith, 1991) and to some extent in listening and speaking in L1 (see for example Fisher & Scriven, 1997). The linguistic indicators of critical thinking have also been investigated in writing in the second language (see for example Stapleton, 2002), as well as in online discussions (Mackey, 1977; Uzuner, 2007).

Speech and writing have considerably different characteristics (McCarthy & Carter, 1994). Brown and Yule (1983) explain these differences in detail. According to them, speakers have the advantage of voice, facial, postural and gestural expressions and they can observe the listener and modify what they say, but are under more demanding conditions than in writing such as monitoring what they say

and making sure that it matches with their intentions, while at the same time planning their next utterance and fitting it into the overall context, as well as monitoring reception by the listener. They also have no record of what has been said before and no notes for what they want to say next. They are under pressure to keep talking and have to publicly correct themselves if the words that they utter are not what they intend to say. In addition, they may suffer from exposing their feelings and having to speak clearly and make immediate responses. Writers, on the other hand, may go over what they have written, pause with no fear of being interrupted, take their time in choosing words, check the process they are going through, reorder what they have written or even change their mind and cross out things they have written privately. In spite of all these differences between speaking and writing, critical thinking in oral L2 has not been studied in the literature

In Turkey, the EAP programs of English-medium universities aim to improve students' language to an academic discourse level that enables them to reflect their critical thinking, as well as to improve their critical thinking in all skills of literacy. As most of these courses aim to teach students how to express their thoughts more effectively in English, students are generally required to write and speak a lot. However, teachers complain that they cannot see any signs of critical thinking in what students produce. Students also complain that they look like they cannot think because they cannot express their opinions comfortably enough in L2. This is the case especially in oral L2 as there is no time to plan what one is going to say. Thus, the ability of students to express themselves critically in the L2 is under question.

Research Questions

This study aims to address the following research questions:

1. What is the amount of critical thinking expressed in tasks that invite critical thinking in EAP classrooms in Turkish universities?
2. What is the nature of language when vocalizing critical thinking in tasks that invite critical thinking in EAP classrooms in Turkish universities?

Significance of the Study

This study will contribute to the literature by exploring the amount of CT expression in EAP class discussions conducted in L2 and the nature of EAP students' language while expressing CT. The conclusions that will be reached from this study might contribute to the understanding of the oral discourse of CT in L2 and to informing teachers about the factors that might increase the quality of L2 in tasks that involve CT.

The results of this study will provide insights into some Turkish university EAP students' existing ability to express CT in their L2 discourse in class discussion tasks and the spoken discourse of critical thinking skills. In this respect, it will shed light on the discussion of whether these students are able to express their critical thinking orally or not. Therefore, it may be useful in developing institutional strategies to enhance the expression of CT in EAP classes in English medium universities in Turkey by providing a picture of the existing level and nature of oral expression of CT in class discussions.

Conclusion

In this chapter the background of the study, statement of the problem, research questions and significance of the study have been discussed. The next chapter will present the relevant literature on critical thinking. The third chapter presents the methodology and describes the participants, materials, data collection procedures and data analysis procedures of the study. The fourth chapter describes the results of the data analyses. In the final chapter, the findings, pedagogical implications, limitations of the study and suggestions for further research are discussed.

CHAPTER II: LITERATURE REVIEW

Introduction

The aim of this study is to investigate the amount of critical thinking expressed and the nature of language used in class discussions in English for Academic Purpose (EAP) classrooms in Turkish universities. In this chapter the literature in the field will be reviewed. First, the meaning of critical thinking will be reviewed. Next, models to investigate critical thinking in discourse will be investigated. In the subsequent section, its place in education will be covered. The last section will address the relationship between critical thinking and the four language skills.

Defining Critical Thinking

Defining critical thinking (CT) has always been a problematic and controversial issue. Cuban (1984) says that the word "troublesome" is very polite and he thinks that "the area is a conceptual swamp" (p. 676). Lewis and Smith (1993) add that there has been little progress in defining CT since Cuban. Bailin (1998, as cited in Bailin & Siegel, 2003) lists three problems in trying to define CT:

> 1- it is impossible to determine whether particular mental operations correlate with particular cases of good thinking; 2- there is no particular set of procedures that is either necessary or sufficient for CT and 3- terms denoting thinking (for example, classifying, observing, hypothesizing) refer not to mental operations or processes but rather to different tasks requiring thinking. (p. 181)

Having said that, it is still beneficial to look at some classic definitions of CT to get a general idea about what is or what is not CT. First, John Dewey (1910), the pioneer of the modern critical thinking tradition, defined "reflective thinking", which

is another term to refer to critical thinking (Shermis, 1999), as "active, persistent, and careful consideration of a belief or supposed form of knowledge in the light of the grounds which support it and the further conclusions to which it tends" (p. 6). His definition inspired other scholars attempting to define CT, such as Ennis (1989, as cited in Fisher, 2001, p. 4), who defined it as "reasonable and reflective thinking that is focused on deciding what to believe or do".

However, CT definitions remain vague and broad when they do not include what a critical thinker does in behavioral terms. Therefore, Newman's definition is more concrete and clear as he differentiates between critical and non-critical thinking based on classroom observations and interviews. To him, CT "challenges students to interpret, analyze or manipulate information" (1990, as cited in Lewis & Smith, 1993, p. 133) whereas non- critical thinking skills necessitate mechanically applying previously memorized information only, i.e., applying formulas. Barnet and Bedau's (2002a) definition also presents concretely what critical thinking involves. They define CT as eliminating fantasies and simple judgments and consciously searching for hidden assumptions, and different aspects and elements.

Richard Paul's contemporary definition differs from others. He states that CT is thinking about and improving the quality of one's own thinking (1993). Paul's definition differs from those previously given in that he emphasizes metacognition, or thinking about thinking (Fisher, 2001).

In 1990, in a project of the American Philosophical Association, a person engaged in CT was described as using a core set of cognitive skills of analysis, interpretation, inference, explanation, evaluation, and self-regulation to form a

judgment and to monitor and improve the quality of that judgment (as cited in Facione, Sanchez, Facione, & Gainen, 1995).

Among these definitions of CT, Ennis's definition will be used as a guiding statement, as it is neither too limited to only thinking about thinking, as Paul's definition is, nor too broad to give any kind of boundaries.

Models Designed to Investigate Critical Thinking in Discourse

Investigating CT indicators in discourse is useful to exemplify what is meant by CT, as definitions may remain abstract no matter how detailed they are. For this reason, CT models, which present and describe different dimensions, components, behaviors or mental operations of CT, have always been popular in the literature since they define CT more concretely and precisely than definitions. Many scholars (see for example B. S. Bloom, Thomas, & Madaus, 1971; Kamin, O'Sullivan, Younger, & Deterding, 2001; Newman, Webb, & Cochrane, 1995) have designed different models for indicators of CT. For the purposes of relevance, only those which have been used or which are suitable for use with an analysis of spoken discourse will be described here. These models mainly fall into two categories: theory-driven, which all originate from an L1 perspective, and research-driven, some of which have been designed for L2.

Theory-Driven Models of CT

General Models

Perhaps one of the most well-known models of CT is that of Bloom (B. S. Bloom, Thomas, & Madaus, 1971). The reason for this may be that his model was rather specific to educational contexts. His taxonomical model of educational goals of CT included knowledge, comprehension, application, analysis, synthesis and evaluation. The first three are considered to be lower order thinking skills and the last three are considered to be higher order thinking skills. Knowledge objectives, according to Bloom et al., imply "recall or recognition of specific elements in a subject area" (p. 141). Comprehension objectives are described with the help of three hierarchical operations: translation, interpretation and extrapolation, from the lowest to the highest. Application is "the ability to apply principles and generalizations to new problems and situations" (p. 155). The fourth operation, analysis, is about breaking a problem or a communication into its parts so that the relationship between ideas, the arrangement and the organization are clearer. The next type of objectives in Bloom et al.'s list is synthesis. It involves putting together the parts to create a whole that was not there before. This step is considered a type of divergent thinking, which is believed to be necessary for creativity. The highest operation in Bloom's taxonomy is evaluation. It is defined as making judgments about the value of a material, idea and so on by using a criteria or standards. Each step in this taxonomy involves those that come before it. As he also illustrates each category with what students do in the classroom environment, his taxonomy is considered to be very helpful to educators.

Garrison (1992, as cited in Newman, Webb, & Cochrane, 1995) offers an in-depth CT model to be used when students are engaged in problem solving. He identifies five stages of CT. These stages are problem identification, problem definition, problem exploration, problem evaluation/applicability and problem integration. In problem identification, learners examine a problem for a basic understanding, after a trigger event. In problem definition, learners analyze the situation in order to understand what is not apparent in it, i.e., values and beliefs which underlie its statement, using experience to form an approach to its solution. Problem exploration is the stage in which learners get insights on the problem and elaborate on the possible solutions. In problem evaluation, learners judge the solutions and ideas and make a decision. Finally, in problem integration, learners implement their decision. As this model is for problem solving, it is different from Bloom's model in that it gives a process or route for learners to follow.

Speaking

Fisher and Scriven (1997) have designed a model for the productive skills that does not seem to have stages through which students should progress as other theory-driven models do. Rather, they look at CT holistically when they provide us with a set of criteria that critical communications, writing and speaking, should meet: "veracity of claims and assumptions, soundness of inferences made or implied, suitability to audience, clarity of the presentation, comprehensiveness to the extent appropriate, conciseness to the extent appropriate, originality and power of the presentation" (pp. 103-104). Fisher and Scriven seem to have looked at what students produce in terms of CT rather than the process they go through while solving a problem.

Research-Driven Models of CT

Models Designed to Measure CT in L1 Speaking

Several models to measure critical thinking in L1 discussions have emerged from research studies. Among these are Kamin, Deterding and O'Sullivan (1998). They investigated the amount and types of CT in third year medical students' face-to-face discussions of cases. However, their comprehensive model might be useful for other fields of study as their categories are not specific to medical cases. Only their language samples are field specific. They arrived at this model by collating categories of CT from the literature but its main categories were borrowed from Garrison (1991, as cited in Kamin, O'Sullivan, Younger, & Deterding, 2001). By adding categories from the literature and omitting others when necessary, the researchers arrived at mutually exclusive categories. Thus, one can say that this model emerged as they were conducting the discourse analysis. Their model is as follows:

1. problem identification (imparting new information)
2. Problem description (clarifying or agreeing on terms and concepts, bringing outside knowledge to bear on a problem)
3. Problem exploration (linking ideas, interpretation, justifying)
4. Applicability (applying practical utility)
5. Integration (teaching each other, offering critical assessment, group process issues)

As well as having the same categories, both Kamin et al.'s (2001) and Garrison's (1991, as cited in Kamin, O'Sullivan, Younger, & Deterding, 2001) studies were intended for use in analyzing problem-solving situations.

Models Designed to Measure CT in L1 Online Discussions

Online discussions bear elements both from speaking and writing discourse. Indeed, if we thought of speaking and writing as the two ends of a continuum, online discussions would fall somewhere closer to speaking than to writing on this continuum. Kern (1995) finds their light, informal style, direct interpersonal address, fast topic shifts, and recurrent deviations from the subject just akin to spoken discourse. To this end, making use of the models developed for online discussions will be very helpful to analyze spoken discourse in the absence of models designed specifically for spoken discourse.

A detailed model designed to measure CT in online discussions was developed by Newman, Webb and Cochrane (1995), to serve the purpose of comparing online and face to face discussions. Like Kamin et al., they combined several other theory-driven models. In their model, all categories include descriptors of behaviors that indicate CT and non-CT. Their categories are relevance, importance, novelty (new information, ideas, solutions), bringing outside knowledge/experience to bear on the problem, dealing with ambiguities (clarified or confused), linking ideas/interpretation, justification, critical assessment, practical utility (grounding) and width of understanding (complete picture). An example of a pair of CT/non-CT descriptors in the category of novelty is “new problem-related information” versus “repeating what has been said” respectively. Their model was shown to be capable of sensitively gauging differences in CT in their research. However, the category of importance seems a little vague and subjective, as it requires only a judgment of whether points raised in discourse are important or unimportant.

Another model was developed by McLean (2005). She analyzed CT in student messages posted during computer conferences. She started with a combination of categories from theory-driven models, and then made adjustments to her model while using it to analyze the discourse of the posts. The first dimension of her model addresses the types of CT and it includes two levels, with the second level describing the expected behaviors to help the analyst identify in which first-level category the student is operating. The categories are clarification of the thesis, problem or question, making inferences and interpretations, supporting inferences and interpretations, and making value judgments. The second dimension concerns the quality of CT and it includes clarity, relevance, depth, logic, preciseness, breadth, and support elements. This dimension uses the elements of CT as stated by Paul (2007). Having these two dimensions gives the model a certain precision, by clarifying exactly what to measure.

Models Designed to Measure CT in L2 Online Discussions

Uzuner (2007) designed a CT-indications model specifically for online discussions by ESL students. She investigated CT in two different categories, educationally valuable talk and educationally less valuable talk, and exemplified each category with specific language uses. Most of her categories of educationally valuable talk are very much compatible with CT and therefore, they are worth mentioning here. These are exploratory, argumentational, critical, reflective, interpretive, analytical, informative and implicative talk. However, not all educationally valuable talk necessarily entails CT. For example, the "invitational" category is educationally valuable because learners are inviting others into the discussion, which shows that all members are getting a chance to talk, the group

dynamics are working well and perhaps there is sharing of knowledge, which is preliminary to critical knowledge-building in a group discussion, but this category is not directly related to CT. The categories of educationally less valuable talk scheme are affective, judgmental, experiential and reproductional and miscellaneous. Her model clearly defines what educationally valuable talk is in the light of CT, which makes it a practical guide to assess the value of online discussions. It can also be said that it could provide some valuable categories for a face-to-face discussion criteria.

As can be seen, the common point among most of the models of CT in the literature is that they started with established categories and were then revised in the process of discourse analysis (see for example McLean, 2005; Newman, Webb, & Cochrane, 1995). As a starting point for my study, I find the studies by Garrison, (1992, as cited in Newman, Webb, & Cochrane, 1995), Kamin et al.(2001), Mclean (2005), and Newman et al. (1995) the most useful since some of the categories they used are very likely to be seen in spoken discourse, too.

Critical Thinking and Education

The difficulty in defining critical thinking does not mean that it should not be considered as an educational ideal. Asking why CT should be in education is like asking why there should be education at all (Norris, 1985). Bailin and Siegel (2003) express four reasons for fostering critical thinking in students. First, students should be treated as persons with respect. Thus, they should be treated as individuals who are capable of deciding for themselves as to what to do and what to believe. Therefore, teachers should help them to develop the skills to judge for themselves. Second, education prepares them for adulthood and this cannot be achieved by imposing pre-determined roles on them. They should be self-sufficient and self-

directed adults, who can think critically. Third, rational traditions that are at the heart of education, i.e., mathematics, science, literature, art and so forth, have always depended upon and required CT. Last, democracy requires CT from its citizenry as it relies upon good reasoning about issues such as politics, media and so on.

Two concepts stand out among methods to enhance CT in education: collaborative learning and problem solving. There is agreement among scholars that collaborative learning enhances critical thinking skills (see for example Anderson, Howe, Soden, Halliday, & Low, 2001; Gokhale, 1995; Klein, 1993; Newman, Webb, & Cochrane, 1995; Schamber & Mahoney, 2006; Windschitl, 2002). Collaborative learning is an instructional method in which students work in a group towards a common goal (Gokhale, 1995). Research clearly shows that students' critical thinking skills benefit from shared learning environments. Gokhale (1995), for example, examined the effectiveness of individual versus collaborative learning in tasks requiring CT with engineering students and found that students who participated in collaborative activities did significantly better on critical thinking tests than students who worked individually. An important point one should keep in mind, however, is that clear group goals are important to enhance the quality of CT and decision making of groups (Schamber & Mahoney, 2006).

Problem solving and problem based learning, on the other hand, are also usually associated with critical thinking (Kamin, O'Sullivan, Younger, & Deterding, 2001). Problem based-learning is an approach to education in which the starting point of learning is a problem for which there is not a ready answer, the knowledge that students are expected to learn is organized around problems and not disciplines, and students are given responsibility for their own learning (Bridges, 1993). CT is a

larger concept that is expected to be enhanced with problem-based learning (Kamin, O'Sullivan, Younger, & Deterding, 2001).

CT and Language Learning

Many would agree that rote learning is especially ineffective in language learning. The reasons for this ineffectiveness are, first, rote learning is a mechanical process which cannot encompass all the complexities of human learning and what characterizes human learning is meaningful learning. Second, it is long-term retention that determines learning. Yet, long-term retention cannot be achieved through rote learning as it is vulnerable to interference, which leads to forgetting (H. D. Brown, 1972).

Moreover, being a critical thinker and being a successful language learner are usually associated with each other. For example, Carroll (1977, as cited in Vermillion, 1997) states that good language learners are those who try to understand the logic of native speakers while using the language. In addition, Rubin (1979) concludes from her research that good language learners can make good guesses and inferences, which are considered to be critical thinking skills.

However, opposing views exist. Atkinson (1997), in her commonly cited article, asserts four reasons to think carefully before implementing critical thinking in language classrooms too enthusiastically. First, what we refer to as critical thinking might just be common sense as a social practice rather than a teachable concept. Second, too much emphasis on critical thinking might marginalize alternative methods of thought. Third, it might be a culture-specific phenomenon and some cultures might even be transferring the opposite way of thinking to their children.

Thus, teaching critical thinking in the language classrooms might be “less straightforward than we assume” (p. 72). Last, the research results are much too complicated to indicate that the thinking skills taught in the classroom are transferable to contexts outside the classroom.

CT and English for Academic Purposes

Whether or not CT is deemed appropriate in general language learning, developing students’ CT skills is of particular importance in EAP courses. The fact that CT is being increasingly discussed in higher education provokes discussions in language teaching. These discussions affect EAP since its main purpose is to give students all the necessary skills to succeed in their university careers (Evers, 2007). In Western cultures, it appears that being critical and analytical is vital for success in university study. Thus, the transition problems faced by non-native speakers are usually considered to be culture-related and not language-related. Students are not sure of what is expected of them in readings, lectures and assignments at university as these expectations are not explicitly stated (Elsegood, n. d.) Therefore, bridging this gap to address the first year transition problems has become a priority for EAP.

The general acceptance of the need to teach CT is naturally reflected in the curricula of EAP courses at many universities. For example, at Middle East Technical University, in Ankara, Turkey, the Modern Languages Department, which offers EAP courses to all freshmen and sophomores, aims to develop the CT skills of students explicitly in all the goals in its curriculum policy document (*Middle East Technical University School of Foreign Languages Curriculum Renewal Project*, 2004-2005). The EAP programs of the University of Prince Edward Island, Canada ("EAP at the University of Prince Edward Island", 2003-2004), Miami Dade College,

Miami ("EAP 1121-1620 reading competencies, levels 1-6 ", 2000) and University of Calgary, Qatar (Brookfield, 1987) are a few other examples from the world.

Critical Thinking and the Four Skills of Language

More specifically in language learning, there is also a relationship between CT and the four skills of language. Much has been said in the literature about the relationship between CT and reading, writing, listening and speaking. Not only can one find scholars who study these skills alone, i.e., the relation between CT and only reading or only writing and so on (see for example Neilsen, 1989; Risinger, 1987), one can also find studies that analyze these skills' combined relationship with CT as these skills are closely interrelated (see for example Fisher & Scriven, 1997; Paul, 1990). The following sections will describe some of these relationships.

CT and Reading

To start with, the relationship between reading and CT has been discussed the most among the language skills in the literature. The importance of critical reading comes from its being seen as a way to gain knowledge, and knowledge means power. Critical thinking and reading abilities are essential for participation in society because they are what give people their independence of thought and action. Therefore, people who lack critical reading skills will be ignorant and rely on others' thoughts (Neilsen, 1989).

Having mentioned the importance of CT in reading, defining what constitutes critical reading is important. For this purpose, it would be useful to view what kinds of sub-skills and applications the scholars who write on critical reading focus on. It can be seen that many scholars concentrate on many overlapping sub-skills. Rasool,

Banks and McCarthy (1993) provide learners with critical reading training through exercises on identifying underlying assumptions, understanding facts and interpretations, generating reasoned conclusions from facts, exploring writers' perspectives, building connections, reasoning deductively and inductively, understanding descriptive and prescriptive arguments and counterarguments, analyzing and critiquing arguments, understanding errors in reasoning, creating strategies for solving problems, recognizing and locating research, and evaluating evidence. While incorporating most of these components, Flemming's (2000) view of critical reading skills also includes using contextual clues, outlining, understanding implied main ideas, synthesizing, inferring, distinguishing relevant from irrelevant, identifying the purpose and the tone of the writer, understanding figurative language and recognizing bias. Smith (1991) adds understanding the connotations of words to this list of critical reading skills.

Another important aspect of critical reading relates to questions asked by both the teacher and the students. Regarding teachers' questions, Smith (1991) states that critical thinking does not happen automatically and teachers should evoke an inquiring attitude in the classroom and provide students with ways to think critically. Therefore, teachers' questions about readings should not only require students to restate what is in the reading passage but also to analyze and evaluate it. Smith divides restatement questions into three groups: detail questions, sequence questions and main idea questions (p.10). According to him, these kinds of questions are necessary to check that main concepts are understood but not deep enough to make the experience of reading rich. They are the first set of questions to be asked. A second set of questions is analysis questions, such as "In what season of the year

might this story have taken place? What makes you think so?" or "How did the author really feel about the animals? How could you tell?", which require analysis and inference to go beyond what is stated. The other set of questions that Smith (1991, p. 11) presents is evaluation questions, which require students to distinguish facts from opinions, to identify assumptions, and to judge the credibility of the author. These questions are necessary to prevent students from seeing the textbook or the author as the most powerful authority figure, and they inject controversy and debate into the classroom. Examples include: "Was Billy's reaction the best one under the circumstances? Why or why not? " or "How does this author know so much about forest rangers?".

The second and perhaps more important type of questions is students' questions, which they ask of themselves or of the reading text. Smith (1991, pp. 11,12) adds that a consequence and purpose of teachers' questions should be to train learners to generate their own questions when reading instead of waiting for the teacher. Students should be self-stimulated critical readers, who ask questions that seek information, summary, causal relationships and evaluations and who aim to see beyond what is obvious. It is possible that the types of questions that Smith recommends teachers and learners to ask will help learners to develop the aforementioned sub-skills of critical reading mentioned by Rasool, Banks and McCarthy (1993), Flemming (2000) and Smith (1991) himself.

In L2 reading the discussion seems to be around whether English teachers should be engaged in teaching critical thinking skills to students and whether critical reading can be taught (Day, 2005). Wallace (2003) asserts that reading is a public and social act. She claims that critical reading is relevant to foreign language

learning and teaching in two ways. First, reading "allows students to draw more fully on their existing linguistic resources and to stretch them at the same time" (p. 199). Second, as students seek clarity and precision, grammatical accuracy may be an outcome.

CT and Writing

Writing is another skill which is seen to be directly related to critical thinking. Risinger (1987) points out that if designed appropriately, writing assignments are one of the most effective tools to improve CT. He states that four different types of writing are the most effective to enhance CT skills: reporting, exposition, narration and argumentation. These types of writing can foster CT skills by activating background knowledge, teaching to differentiate between relevant and irrelevant information, paying explicit attention to metacognitive processes and creating a culture of questioning (Ruland-Parker, 1999).

It is important to see at this point that there is a mutual and interactive relationship between reading and writing skills. First, all the skills necessary for critical reading have equivalences in writing as both require recognizing supporting, contradicting, vague, clear, false, insightful, prejudiced or conforming ideas and actively reconstructing meanings (Paul, 1990). Another relationship between writing and reading is that writing can be used to develop critical reading skills as it requires learners to verbalize, monitor and revise their own understanding of the texts they read. For example, when the teacher asks the students to freewrite or ask questions about the heading of the reading passage before reading or to summarize or to reflect on the passage after reading, writing turns the reading process into an activity of constructing meaning (Stahl, 1991).

In L2 writing, the central discussion seems to be different. It centers around whether the idea that ESL learners cannot think critically is just a prejudice or a fact. Stapleton (2002), in his study that is commonly referred to by other scholars in the literature, analyzes Japanese students' writings. He finds evidence that, contrary to the common belief about Asian learners, Japanese learners can object to authority and show elements of critical thinking when writing in English. He relates this to the current trends in the Japanese education system. In another recent study Alagozlu (2007) analyzed argumentative essays by Turkish students using Stapleton's criteria for clues as to the elements of critical thinking (claims, kinds of reasoning, the extent of evidence, recognition of opposing arguments and refutation, and fallacies) and individual voice. The results show that the students do not usually support the claims they make with sufficient evidence or reasons from the texts they read, or with sensible and relevant conclusions, and contradictory arguments in the texts do not seem to be recognized and refuted. The researcher relates this to the traditional education system in Turkey and concludes that EFL students need to be supported in terms of critical thinking skills to overcome difficulties in writing.

CT and Listening

Listening can also be considered to be parallel to reading and writing, in that critical readers and writers can also listen critically as the challenge is the same. Still, critical listening can be more difficult for a student because there is no chance to go back and listen again (Paul, 1990).

Some scholars consider reading and listening skills together. Fisher and Scriven (1997, p. 97), for example, call this pair "critical observing" as they are receptive skills by nature. They describe critical observing as having four levels. All

the levels include mastering the specific skills mentioned previously as components of critical reading. The first level is active understanding, which includes reading or listening between the lines, outlining, summarizing, paraphrasing and identifying generalizations, emotional statements, facts and the writer's position. Level two is active inquiry and it includes identifying and creating comparisons, going to secondary sources of the same author, asking someone else, reformulating and adding to the topic for points that are left open-ended in the text. The third level is active generalization, which is about testing the trustworthiness of the arguments with experience, knowledge or sources. This step is considered to be *moving* to the metalevel. The last step of critical observing, as in reading and listening, is active self inspection, in which the interpreter thinks about the weaknesses, such as biases and limitations, and the strengths in his/her own thinking.

CT and Speaking

There is not much in the literature about CT and speaking. The scholars who write about spoken CT mostly refer to it as *language*, which means they consider all the productive skills together, i.e., speaking and writing. Fisher and Scriven (1997), for example, consider writing and speaking skills together as they are both productive skills, add presenting to these skills, and call this set of skills "critical communication" (p. 101). They state that these competencies are a part of the critical thinking process for several reasons. First, these skills include a self-critical process and can be subject to critical interpretation. Moreover, critical communication includes self assessment and improvements regardless of the existence of the audience. Perhaps more importantly, the writers state that communication is a part of argument because we must communicate with people whenever we are interested in

making a claim, convincing others, explaining our position and so on. Therefore, critical communication must meet the same criteria as critical thinking, i.e. validity, concision, clarity, power and so on, which suggests that critical communication is a mixture of "analytic interpretation and effective communication" (p. 102). Another reason that CT includes communication is that people usually benefit from brainstorming with others when they are not sure about their position about a subject, in which case communication becomes an internal segment of the critical thinking process. Yet another reason Fisher and Scriven present is that oral argumentation involves both critical listening and speaking, with neither being less important than the other. Moreover, we never say that someone is a good critical thinker but cannot express critical thought in his/her native language. Therefore, "expression is part of the thinking" (p.102). As a final point, Fisher and Scriven add that communication is not something we do after we have finished the thinking; thinking and expression occur concurrently. Thus, critical communication refers to "skilled and active critical review" (p.102) and it is a part of CT.

When talking about critical oral communication, one should also consider the specific language to be used. When students are reasoning, if they are able to use analytical vocabulary such *as infer, conclude, criteria, point of view, relevance, issue, elaborate,* they can make their thinking more accurate and clearer (Paul, Binker, Martin, Vetrano, & Kreklau, 1989).

Research on CT and Speaking

Research on CT and speaking has mostly been done to evaluate CT in group discussions, as there is a clear link between CT and interaction (see for example Gokhale, 1995). A study conducted by Kamin, O'Sullivan, Younger and Deterding (2001) analyzed the CT discourse of medical students in group work discussions in problem-based tasks. They videotaped the students' discussion sessions and coded the CT in the discussions with a framework they designed. Two groups of students were used and one of them was given the problem in text format, while the other one watched the same problem from a video-recording. The researchers showed that different levels of CT in problem based situations can be coded sensitively using their criteria. Newman, Webb and Cochrane (1995) also measured CT but unlike Kamin et al. (2001), they compared face to face and computer-based environments. They assessed how much and what type of critical thinking is done in both situations with the framework they designed based on other frameworks. They also found that their framework sensitively measures the differences in CT, using discourse analysis. The details of their findings were discussed in a later paper (Newman & Johnson, 1997). The face-to-face discussions stimulated more creative problem exploration and idea generation and the computer-supported discussions generated better linking ideas, interpretation and problem integration.

In brief, being able to use critical thinking skills is very important in language learning, as well as in all aspects of education. Additionally, Turkish students need a lot of support in that (see for example Alagözlü, 2007). However, although the relationship between CT and reading, writing and listening has been discussed in the literature, the spoken language of EFL students has not been investigated in terms of

CT. Likewise, the nature of the language of CT in oral discourse has not been investigated.

Conclusion

In this chapter, the literature on critical thinking was reviewed. The meaning of critical thinking, models to investigate it in discourse, its place in education and the relationship between critical thinking and the four language skills were discussed. It has been revealed that although the relationship between CT and reading, writing and listening has been discussed in the literature, the spoken language of EFL students has not been investigated in terms of CT, and neither has the nature of the language of CT in oral discourse. The study described in the next chapter attempts to fill in the gap in the literature by investigating the amount of CT expression in tasks which invite CT and the language nature while doing such tasks in EAP classrooms in Turkish universities. In the next chapter, the methodology used in this study, including participants, instruments and data collection and analysis procedures, will be covered.

CHAPTER III: METHODOLOGY

Introduction

This study was designed to investigate the ability of Turkish EAP students to express CT in their oral L2 discourse in discussion tasks. It attempted to address the following research questions:

1. What is the amount of critical thinking expressed in tasks that invite critical thinking in EAP classrooms in Turkish universities?
2. What is the nature of language when vocalizing critical thinking in tasks that invite critical thinking in EAP classrooms in Turkish universities?

This methodology chapter is composed of four parts. In the first part, the participants in the study will be described. In the second part, the materials and instruments used will be explained. The third part will present the data collection procedures. In the last part, information on how the data was analyzed will be given.

Participants

The study was conducted at Middle East Technical University (METU) since the EAP program at METU has CT among its curricular goals (*Middle East Technical University School of Foreign Languages Curriculum Renewal Project*, 2004-2005). The participants were two classes of first year METU students who were taking the ENG102 (English for Academic Purposes II) course. The Eng 102 course was chosen because it is a theme-based course. In this way, a task related to a course topic could be employed, thus ensuring a similar amount of background

knowledge among the participants. Having two classes increased the reliability of the results.

There were 34 students who were all between the ages of 17- 21, in two classes. Twenty four of these students were in one class and ten of them were in the other class. There were five females and 29 males. Twelve students had not been to the preparatory school at university. One student was not taking this course for the first time. Their level of English was upper-intermediate to advanced. Twenty-two students were studying Electrical and Electronic Engineering. Nine students were majoring in Computer Engineering, one in Industrial Engineering, one in Metallurgical Engineering and one in Physics Education. The teachers of the two classes were different.

Instruments

Materials used in this study included a discussion task, categories of critical thinking to code the discussion, and a set of criteria to examine the nature of the language in the discussion. Below, these instruments will be described in detail.

The task

The researcher determined that the task should not require the participants to do much reading as then it would also be testing their L2 reading comprehension or critical L2 reading skills. Having said that, the students might have read about the topic before the actual research task since it is a theme-based course. This was an expected situation and it also helped to assume a minimum common amount of content knowledge of the students. The important point was that the task that students would do should not include heavy reading.

Having taken all these points into consideration, the task was chosen. The task was a whole class discussion task about invasion of privacy, which was the topic of the unit students were covering. The students were divided into three groups. One group represented celebrities, another one the paparazzi and the other group, the public. The students were assigned to groups randomly and given 10 minutes to prepare for the discussion as a group. The class teacher led the discussion by nominating the students who raised their hands to take a turn.

The framework for the categories of critical thinking

In order to determine the amount of expression of CT in the discussion, it was necessary to collect categories of CT from the literature in accordance with the categories of CT in the discussions. Some other categories that were not in the literature have also emerged from the analysis of the discourse of the discussions. The arguments put forward in the discussion should be supported by thought in one or more of these categories. The categories are shown below.

Critical Thinking

1. Clarifying/defining (see for example Kamin, O'Sullivan, Younger, & Deterding, 2001; Garrison, 1992, as cited in Newman, Webb, & Cochrane, 1995; Uzuner, 2007), (examples, defining terms, pointing at another aspect of the issue, metaphors)
2. Analysis/synthesis (B. S. Bloom, Thomas, & Madaus, 1971), (similarities, differences)
3. Enhanced rephrasing, (with additions)

4. Offering solutions/direction/ a course of action, (see for example Eckberg, 1977, as cited in Bailin, 1998; Garrison, 1992, as cited in Newman, Webb, & Cochrane, 1995) (statements of should)
5. Inference/interpretation (see for example McLean, 2005; Newman, Webb, & Cochrane, 1995), (reasoning, consequential)
6. Brief and triggering arguments or questions (Garrison, 2001, as cited in Meyer, 2003)

Not Critical Thinking

1. Repetition without really adding anything new (Uzuner, 2007)
2. Unclear/unfocused or irrelevant idea or examples (Uzuner, 2007)
3. Accusational or defensive statements or questions without satisfactory explanation or evidence
4. Emotional statements that include feelings (Uzuner, 2007)
5. Logical fallacies (Stapleton, 2001)

The criteria to examine the nature of the language used

After examining many spoken language criteria in the literature, including the Massachusetts Speaking Assessment Criteria (Carter & Nunan, 2002), Student Oral Language Observation Matrix (Bailin, Case, Coombs, & Daniels, 1999) and TOEFL speaking criteria (Atkinson & Ramanathan, 1995), the researcher decided to devise her own criteria as she wanted something simple, easy to use and applicable for transcriptions of oral language. This set of criteria had three components: vocabulary, grammar and effectiveness in conveying the message. This was because these three were the common points of many different sets of criteria that the researcher examined, excluding the aural components of spoken language such as

pronunciation or fluency. Each component was graded on a scale of 1 to 5. A one for vocabulary means “very simple and limited vocabulary, frequent wrong use of vocabulary, hinders understanding”, whereas a five indicates “rich, diverse, correct, extensive use of vocabulary”. A one for grammar is using “basic grammar, simple sentence structures, mistakes, frequent errors, hinders understanding” and a five is “a balance of simple and advanced sentence structures used correctly, able to explain precise points or subtle differences in terms of ideas with the help of grammar”. In the last category, effectiveness in conveying the message, a one indicates “ineffective use of language, very difficult to understand the idea without a lot of interpretation”, while a five means “excellent use of language, ideas are very clearly stated with the help of the language” (See Appendix A for the complete set of criteria for the nature of language used).

Procedure

In the first week of February, 2008, the researcher contacted the head of the Modern Languages Department at METU to ask permission to video-tape two English 102 classes, and to use the departments’ camera and the meeting room. Upon getting permission, an application was made to the ethics committee of METU, stating the researcher’s purpose for using human subjects. Copies of the application documents can be seen in Appendices B and C.

Piloting was done in the third week of February, 2008 with a small group of different students. The piloting was valuable in that it showed the researcher what kinds of ideas might emerge and what kinds of problems might arise.

In the fourth week of February, 2008 and the first week of March, the researcher asked two teachers to carry out the task in their classes. The researcher explained the procedure and the task to the teachers. On the day of the study, the researcher first introduced herself and what she was doing to the class without going into much detail and had students sign the informed consent form. A copy of the informed consent form can be seen in Appendix D. After video-recording the discussions, the researcher distributed the debriefing form to the students. A copy of the debriefing form can be seen in Appendix E.

The discussions recorded were first transcribed and divided into utterances or chunks of complete meaning as in Newman et al.'s (1995) and McLean's (2005) studies. The utterances sometimes lasted for one sentence only, but sometimes they were spread over a few turns in the discussion. Then, the discourse of the transcriptions was analyzed.

To answer the first research question about the amount of CT expression, the transcriptions were coded, first using the categories of CT that were collected from the literature. The categories were modified and extended as the analysis of the discourse was being carried out, so that the categories emerged simultaneously with the analysis. Another rater, who is a native speaker English teacher, was also asked to label the utterances using the final version of the categories, for reliability purposes. However, great difficulty in reaching an agreement on the categories of critical thinking in the utterances was experienced. This appeared to be a result of the fact that the utterances usually included more than one category of critical thinking, which could not be separated from each other. After several trials, it was decided that the researcher and the second rater should simply try to reach an agreement only on

whether the utterances represented CT or not. In order to do this, it was decided that seeing only one category of critical thinking in an utterance was sufficient to mark it as critical thinking. This kind of procedure exists in the literature, such as in Newman et al.'s (1995) study. During this process, the utterances that were totally unclear were also marked. The overall interrater reliability rate was 91.1 %.

To answer the second question, two other raters assessed the utterances in the transcriptions that had been coded as CT, in terms of language used, using the set of criteria that was developed by the researcher. One of the raters was a native speaker English teacher while the other was a native speaker who is not an English teacher. The raters graded the utterances for vocabulary, grammar, and overall effectiveness in conveying the message, which constituted the *nature of language*. The overall interrater reliability was 93.9 %. The utterances that the raters did not agree on were moderated by a third rater.

Data Analysis

First, the utterances of critical thinking, non-critical thinking and unclear language that were agreed on both by the researcher and by the second rater, and the utterances on which an agreement could not be reached were counted and reported in terms of percentage. Second, the averages of the grades given by the two raters for the vocabulary, grammar, effectiveness in conveying the message and for the overall nature of language used in each utterance of CT were reported in terms of percentage.

Conclusion

In this chapter, the participants, instruments, procedures and data analysis were described. In the next chapter, the data analysis procedures and outcomes will be explained in detail.

CHAPTER IV: DATA ANALYSIS

Introduction

This study was designed to investigate the ability of Turkish EAP students to express CT in their oral L2 discourse in discussion tasks. It attempted to address the following research questions:

1. What is the amount of critical thinking expressed in class tasks that invite critical thinking in EAP classrooms in Turkish universities?
2. What is the nature of language when vocalizing critical thinking in class tasks that invite critical thinking in EAP classrooms in Turkish universities?

This study was conducted with the participation of two classes of EAP students studying at Middle East Technical University. Two classes were video-recorded while engaged in a task that required discussion. The discussion was among three groups of students representing paparazzi, the public and celebrities. The first discussion lasted 35 minutes and had 10 participants. The second discussion was about 60 minutes long and had 24 participants. The recordings were then transcribed and the discourse of the discussions was analyzed.

This chapter presents an analysis of the transcriptions of the discussions in terms of the amount of CT expressed, the nature of L2 in a task that invites CT, other features of the CT utterances, other features of the non-CT utterances, and correlations among the scores. The first section presents the data analysis procedures that were followed while determining the amount of critical thinking expressed in

these discussions, and the results of these procedures. The second section presents the data analysis procedures followed while determining the nature of the language used in the utterances vocalizing critical thinking, and is followed by the results. The third section is related to the correlations among the scores of the criteria used to examine the nature of the language used. The fourth and fifth sections present other results related to the CT and non-CT utterances that were obtained during the data analysis.

The analysis of the discussions

First, the transcriptions of the two discussions were divided into utterances by identifying complete meaning units in the discussion. These meaning chunks appeared in three forms. The first form was one turn by one student, representing one utterance when the student expressed the same idea throughout his/her turn. The second form was a few turns of the same student, representing the same idea in spite of interruptions by other students, if the student continued with the same idea. The last kind of utterance usually appeared when a student's turn lasted a long time. In those cases, the students usually talked about several separate ideas within the same turn, each one representing a different utterance. In this way, 83 utterances were identified in the first discussion, which had 10 participants and lasted 35 minutes, and 131 utterances in the second, which had 24 participants and lasted 60 minutes. In total, there were 214 utterances.

After identifying the utterances, the researcher and a second rater attempted to identify the categories of CT. Although they were able to categorize the utterances individually, it was not possible to reach a high percentage of agreement. For example:

Extract 1:

> I want to deter ... eee … determine something. Public is interested in celebrities, because they see in cinema or movie or TV and they wonder whether people are like in the movie or TV. And ... eee ... actually paparazzis are born from this point.

Or

Extract 2:

> I think ... eee ... I don't agree with my friends. I think that celebrities is a part of life over the history because in ... such that ... in the middle age ... there an aristoc ...Yes. Today is people who is millionaire or ... who is businessman is celebrities or film actress. But only thing that change over the history in early age or before one hundred years, people think the celebrities is the same as the other people and don't wonder the others. But now people people is wonders so much things that is life of the celebrities, and love of the celebrities ... go on.

The researcher thought that Extract 1 was in the *inference/ interpretation* category because the student was reasoning about how the public's interest in celebrities has started. On the other hand, the second rater thought that this utterance was in the *clarifying/ defining* category because the student was clarifying how it all started. In extract 2, the researcher thought that it was *analyzing/ synthesizing* as the student was mainly focusing on similarities and differences of today and the past in terms of privacy, but the second rater thought that it was *inference/ interpretation* as he thought that the student was making an interpretation about the situation by looking at the history. One may also say that it is *clarifying/ defining* as well, as the student was trying to clarify the issue by point out its background. Consequently, although some agreement was achieved after long discussions, it was decided that the real point that the researcher and the second rater did agree upon was that most utterances included more than one category of CT.

In spite of the fact that these categories could be recognized in the utterances clearly and that the literature confirmed that they were CT, differentiating them from one another was not easy in this discourse, as shown above, and the ultimate categories that would be decided on after long discussions would still be open to other possible interpretations. One other solution was to try to arrive at mutually more exclusive categories. However, it was decided that this would also be much more difficult than was initially thought as it would require time and expertise. Therefore, having considered that the categories that were identified by the researcher and by the second rater made sense and could be supported by argument, it was decided to look at only the amount of CT expressed in the discussions. Instead of labeling the utterances in terms of their CT categories, the researcher and the second rater labeled them as CT or non-CT with reference to the framework of types of critical thinking developed by the researcher according to the types of critical thinking in the literature (see Appendix F for the critical thinking framework). For this, it was decided that seeing only one category of critical thinking in the framework in an utterance would be sufficient to mark it as critical thinking (see for example Newman, Webb, & Cochrane, 1995). During this process, the utterances that were totally unclear were also marked (see Appendix G for an example of coded transcripts). The percentage of agreement between the researcher and the second rater was 90.36 % for the first class and 91.6 % for the second class. The overall interrater reliability was 91.1%.

Labeling the utterances

CT utterances

The first category of the analysis of the discourse was the utterances that include CT. In the extracts below, two sample utterances that were marked as critical thinking by both raters are shown.

Extract 3:

> But I want to say something. If you don't want to ... eee .... published in public... don't make these activities in public ... like bus … like eee … in such public areas ...

Extract 4:

> I want to say something about the Angelina Julie and the Brad Pitt's new ... new… new born babies. The photo is taken and than it cost … millions of dollars to a journal ... and ... the … journal ...satıl? ... sold ... eee yani ... come to the best seller that month ... It sold...millions maybe … millions of people read it, watched the ... eee … photos and this has became a gossip of the eee ... month at that time. You can't say that ... eee ... celebrities don't want to be taken in a photo … don't want to be in a photo … because this is an advertorial for them … this is an advertisemen ... advertorial for them. If the photo doesn't taken, the Brad Pitt doesn't ... Maybe Brad Pitt doesn't ... eee ... take a role...millions of this movies ... maybe doesn't be ... wasn't be so successful ... like that.

According to the framework that the researcher prepared, in extract 3, the category of critical thinking could be *inference/ interpretation* or *challenging what has been said with brief but sensible and relevant arguments or questions*. In extract 4, the category might be *clarifying/ defining* since the student was giving an example for his argument, or it could also be *inference/interpretation* because the student was also making some interpretations by drawing conclusions from the situation.

Non-CT utterances

The second category of the analysis of the discourse was the utterances that do not include CT. In the extracts below, two examples from this category are shown.

Extract 5:

> Yea. I against them because they will start … they was start ... they started this situation.

Extract 6:

> Yes. It make me angry.

In extract 5, there is just accusation without explanation or evidence. Thus, it was considered to be in the category of accusational. Extract 6 includes a sample from the emotional category.

Unclear utterances

The third category of the analysis of the discourse included unclear utterances. The two extracts below show examples that could be understood by neither of the raters.

Extract 7:

> People who watching you what must do this privacy section? So …

Extract 8:

> But same days same ...while a TV channels eee … programming a eee .... paparazzi show... eee the other one .... again .... eee ... programming another one, so there were no choice to change another channel to ...

Undecided utterances

The fourth category of the analysis of the discourse included the utterances on which an agreement could not be reached by the researcher and the interrater. Extracts 9 and 10 below, show two samples that were categorized differently by the raters.

Extract 9:

> Yes but … who is public? ... You are public, we are public, I mean.

Extract 10:

> We accept that there is a minority among us who abuses ... who abuse ... that situation. I think like the paparazzis, they ... that ... they also … they are also lack of ethics ... and so we blame them we oo we also blame them.

Extract 9 above was marked as CT by the researcher. The researcher thought that this statement fitted into *challenging what has been said with brief but sensible and relevant further arguments or questions* category in the discussion. However, the second rater did not agree that it represented CT. Extract 10 above was not considered to be CT by the researcher on the grounds that it only has *accusational or defensive statements or questions without satisfactory explanation or evidence,* as she did not see any satisfactory explanation as to why the speaker blames the celebrities who abuse the situation. However, the second rater did not agree with it and thought that it falls in the *clarifying/defining* category of CT.

The amount of CT expressed in the discussions

The results of this classification can be seen in Figures 1, 2, and 3 below.

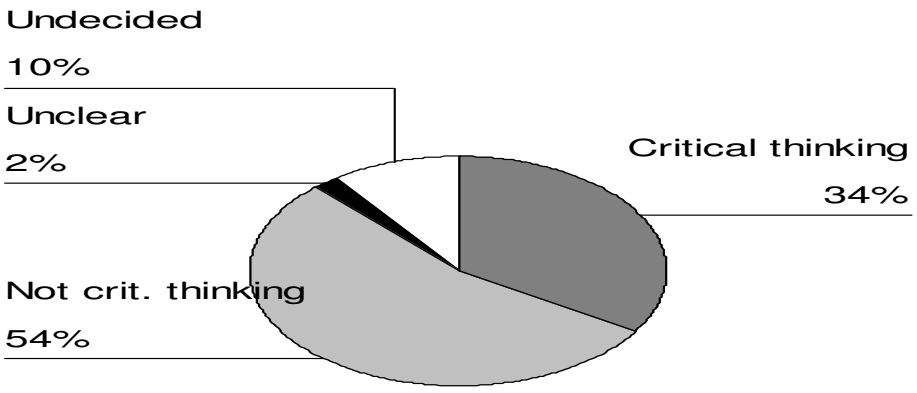

Figure 1 - The amount of CT expressed in the first discussion

In the first discussion, the percentage of non-CT is much more than the percentage of CT. The percentage of non-CT is more than half of the utterances in the discussion. What is more is that, even if all the utterances on which an agreement could not be reached by the researcher and the second rater as to whether or not they included CT had been classified as CT, it would still mean that less than half of the utterances in the class had consisted of CT.

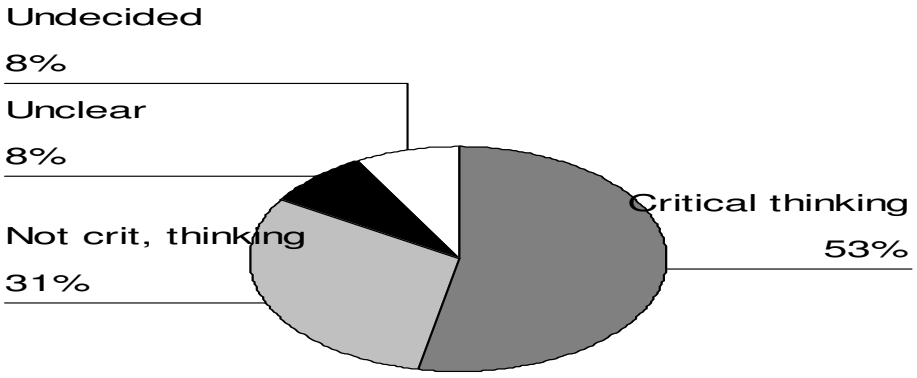

Figure 2 - The amount of CT expressed in the second discussion

Unlike the first discussion, in the second discussion, the percentage of CT expression is much higher than non-CT, and it is higher than the amount of the expression of CT in the first discussion. The percentage of utterances that are unclear is four times more than the first discussion, probably because this discussion was almost twice as long as the first discussion.

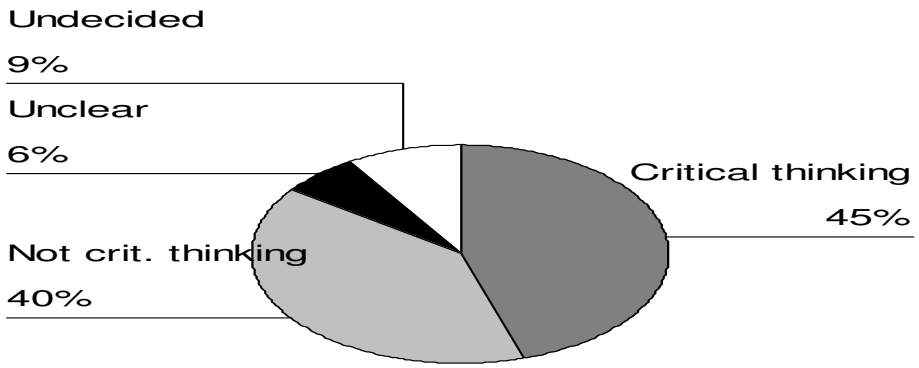

Figure 3 - Total amount of CT expressed in the discussions

In total, the percentages of utterances that include and that do not include CT are not very different from each other although there are slightly more utterances that include CT.

In brief, the percentages of CT expression in the first and the second groups differ considerably in that there is much more expression of CT in the second discussion, which included 24 participants and lasted 60 minutes. Related to this, the percentages of non-CT also differ. In total, the percentages of utterances that include and do not include CT are similar to each other.

Nature of language used in CT utterances

The utterances that were agreed upon as critical thinking by the researcher and the second rater were given to two other raters, both of whom were native speakers of English, one of them being an English teacher. A set of criteria of the nature of the language used that was devised by the researcher was used by the raters (see Appendix A). The raters were expected to rate each utterance in terms of vocabulary, grammar, and the overall effectiveness in conveying the message.

After getting the results from the raters, first, the scores given by the raters which did not differ by more than one point were averaged. The interrater reliability for the different components of the criteria is shown in Table 1 below. Those that differed by more than one point were resolved by a third rater.

Vocabulary	97.9 %
Grammar	98.9 %
Overall effectiveness in conveying the message	89.7 %
Total	93.9 %

Table 1 - Interrater reliability scores for the criteria for the nature of language

Second, the results were analyzed in terms of means, correlations and percentages. Table 2 below shows the means for the three aspects of the criteria and the total scores.

	N	Minimum	Maximum	Mean	Std. Deviation
vocab score	97	2.0	4.5	2.845	.5222
grammar score	97	1.5	5.0	2.716	.6077
Effectiveness in conveying the message score	97	1.0	4.5	2.928	.7217
total score	97	5.0	14.0	8.490	1.4755
Valid N (listwise)	97				

Table 2 - Overall means for each of the aspects of the criterion, and the total score

The means of all the aspects of the set of criteria are not very different from each other, in that all of them are a little higher than half in a scale of 1 to 5, although the slight differences among the means show that these aspects can be ordered from the lowest to the highest as grammar, vocabulary and effectiveness in conveying the message. The mean of the total scores is also similar to the specific aspects of the set of criteria as it is, again, slightly more than half in a scale of 1 to 15.

In the following sections, the results of each component of the set of criteria and the overall score will be presented with examples.

Vocabulary use

The first aspect of the nature of language examined was vocabulary. The utterances were rated on a scale of one to five, one being “very simple and limited vocabulary, frequent wrong use of vocabulary, hinders understanding”, and five being “rich, diverse, correct, extensive use of vocabulary”. The vocabulary scores are as shown in Figure 4 below:

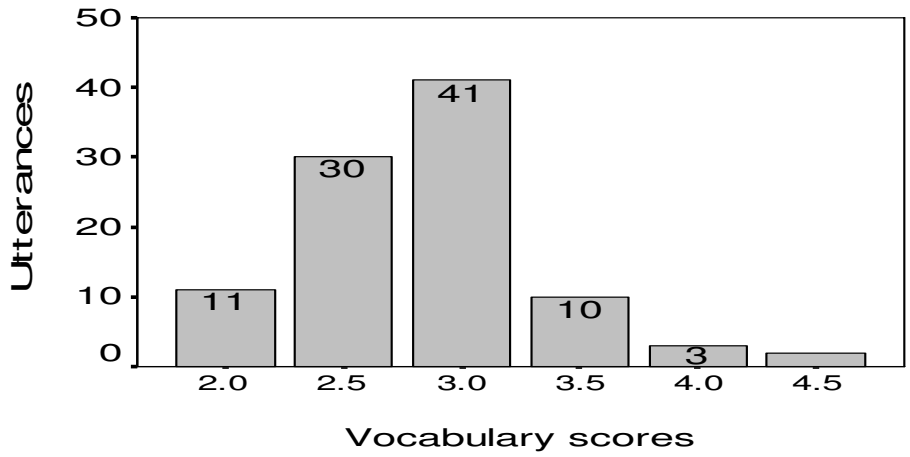

Figure 4 - Vocabulary scores

The vocabulary scores shown in the figure above reveal that the lowest score was 2, given to 11% of the utterances, the highest score was 4.5, representing 2% of the utterances, and the most frequent score was 3, corresponding to 41% of the utterances. An example from each of these three categories will be shown below.

An utterance that was given a score of 2, the lowest score for vocabulary, is shown in extract 11 below.

Extract 11:

> You say … you say ... that the government allows us but when the government is behaving in injustice ways, eee ... don't you have to ee ... correct this wrong issues as journalist? I think, the first issue of journalist is to inform people and eee ... to ... eee … bişeyi açıklamaktır ... yani ...

Perhaps, the most important reason why it was given the lowest score is the use of Turkish, which indicates *very limited vocabulary which hinders understanding* in the criteria. In five of the utterances, students were observed to switch to Turkish when they lacked the necessary vocabulary to explain what was in their minds. Three of these utterances belonged to the same student. In addition, in this extract, there are

mistakes in terms of form and the meaning. The message is conveyed with basic level words.

An utterance that was given a score of 4.5, the highest score for vocabulary, is shown in extract 12 below.

Extract 12:

> YOU should protect your children. You should take them away from the screen and put them into bed.

The reason that it was given a 4.5 could be that all the vocabulary items are used correctly, including phrasal verbs and some topical vocabulary. Although it does not have rich and diverse vocabulary, since there are no other examples of very advanced and academic vocabulary used correctly, this utterance may have been seen to merit a high vocabulary score compared to the other utterances.

In extract 13 below, a typical example of an utterance that was given a 3, the most frequent grade, is shown.

Extract 13:

> But in my opinion, romance is the part of celebrities' lifes and if public interested in their romance and their relationship, that's our job to show that to public.

Characteristically, it includes no or few vocabulary mistakes that hinder understanding but the message is conveyed with very simple and basic vocabulary items, as mentioned in the nature of language criteria.

Grammar

The grammar of the utterances was graded with another scale from one to five, one representing utterances that have "basic grammar, simple sentence structures, mistakes, frequent errors, hinders understanding", and five for "a balance of simple and advanced sentence structures, used correctly, able to explain precise points or subtle differences in terms of ideas with the help of grammar". The grammar scores are as shown in Figure 5 below.

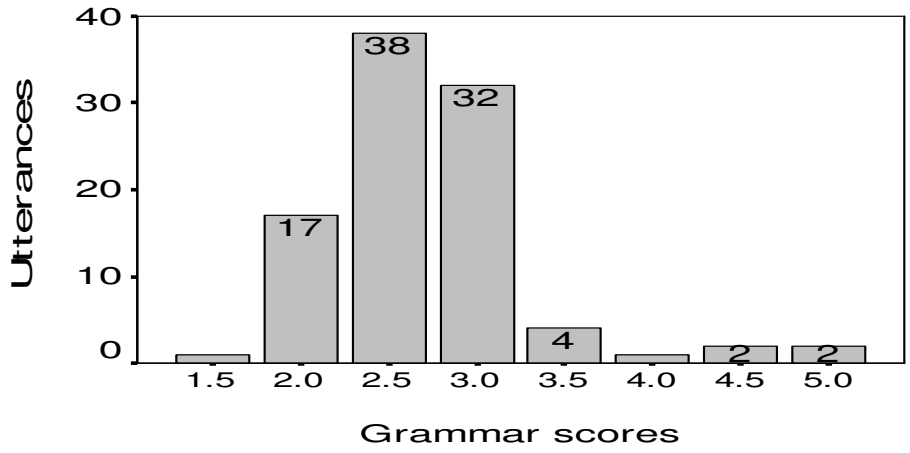

Figure 5 - Grammar scores

The grammar scores shown in the figure above illustrate that the lowest score was 1.5, given to 1% of the utterances, the highest score was 5, representing 2% of the utterances and, the most frequent score was 2.5, standing for 38% of the utterances. An example from each of these three categories will be shown below.

The only utterance that was given a score of 1.5 for grammar, the lowest grade, is shown in extract 14 below.

Extract 14:

> Of course there is people … there are people like ... behaving that but every people ... every celebrity people ... don't like...don't behave that ... if journalists can ... don't ... ilgilenmek? Eğer onlar ilgilenmeselerdi … they don't like ... they don't behave like that.

The reason that the raters rated it so low could be the unfinished sentences and frequently corrected words or structures. Probably, the use of Turkish has also affected the score given for grammar.

An utterance that was given a score of 5 for grammar is shown in extract 15 below, which is the same as extract 12 above, which was given the highest score for vocabulary as well.

Extract 15:

> YOU should protect your children. You should take them away from the screen and put them into bed.

The reason why raters rated this utterance as a 5 could be that it has correctly used structures which are appropriate to the context. Because there are few advanced structures used in the transcriptions, as stated in the grade 5 part of grammar in the nature of language criteria, the raters, most probably, gave this utterance a 5 in comparison to the other utterances.

A typical utterance that was given a score of 2.5, the most frequent grade given for grammar, is shown in extract 16 below.

Extract 16:

> I think ... eee ... I don't agree with my friends. I think that celebrities is a part of life over the history because in ... such that ... in the middle age ... there an aristoc... Yes. Today is people who is millionaire or ... who is businessman is celebrities or film actress. But only thing that change over the history in early age or before one hundred years, people think the celebrities is the same as the other people and don't wonder the others. But now people people is wonders so much things that is life of the celebrities, and love of the celebrities...go on.

Characteristically, it has some grammar mistakes, such as subject-verb agreement or word order, but they do not overly interfere with the message.

Effectiveness in conveying the message

The effectiveness in conveying the message aspect of the set of criteria included a scale from one to five, one being "ineffective use of language, very difficult to understand the idea without a lot of interpretation", and five being "excellent use of language, ideas are very clearly stated with the help of the language". The totally unclear utterances had already been eliminated while deciding on CT and non-CT utterances, and therefore the criteria did not have a 0, which would refer to utterances which were impossible to understand. The scores for effectiveness in conveying the message are as shown in Figure 6 below:

Figure 6 - Scores for effectiveness in conveying the message

The scores for effectiveness in conveying the message shown in the figure above indicate that the lowest score was 1, given to 1% of the utterances, the highest score was 4.5, characterizing 3% of the utterances, and the most frequent score was 3, representing 28% of the utterances. An example from each of these three categories will be shown below.

In extract 17 below, an example of an utterance that was given a 1, the lowest grade for effectiveness in conveying the message, is shown.

Extract 17:

> Let me first make an analogous. The paparazzi group says, claims that public demands us and then we do this job ... Eee ... eee .... But the German public, before the second world war demand Hitler, not demand but they support Hitler to annihilate...annihilate Jewish people. Then, was Hitler arrive to kill Jewish people. Then, I will ... eee …

A 1 in this category means that this utterance requires a great deal of interpretation to be understood. In the extract above, it can be understood that the student is giving an example of Hitler, but in what sense he finds Hitler's time similar to today is not clear at all, regardless of the grammar and the vocabulary of the utterance. This utterance received a 3 for vocabulary and a 2 for grammar.

In extract 18 below, an example of an utterance that was given a score of 4.5, the highest score for effectiveness in conveying the message, is shown.

Extract 18

> We just ... sometimes we just put the photographs, we make no comment and we just leave it to the public.

The utterance is clear-cut in terms of its message in context although it uses simple vocabulary and grammar and has occasional errors. This utterance received a 3 for its vocabulary and grammar.

In extract 19 below, an example utterance that was given a 3, which is the most frequently given grade, is shown.

Extract 19:

> Everyone knows that a singer has a stage personal and an actor has a public personal and a real personality and those photographs that are shot in the public areas does not react ... reflect their real personalities. The all have ...

> eee ... they all ... eee … everyone behaves different in home at home and behaves different in public. So that's just...like role playing ... Even actors, singers, characters, any ... any celebrity, they have public personals, all of them. And it's not real.

In this utterance, the message is clear although it may require some interpretation at times. It has some language errors that may not overly hinder interpretation. It is "pretty understandable" as stated in the nature of language criteria. Its grammar score is 2.5 and its vocabulary score is 3.

Total scores for nature of language

After dealing with each aspect of the criteria individually, the scores for the three components were added together to arrive at a total score. The total scores for the first and the second discussions and in total are shown in figures 7, 8 and 9 below.

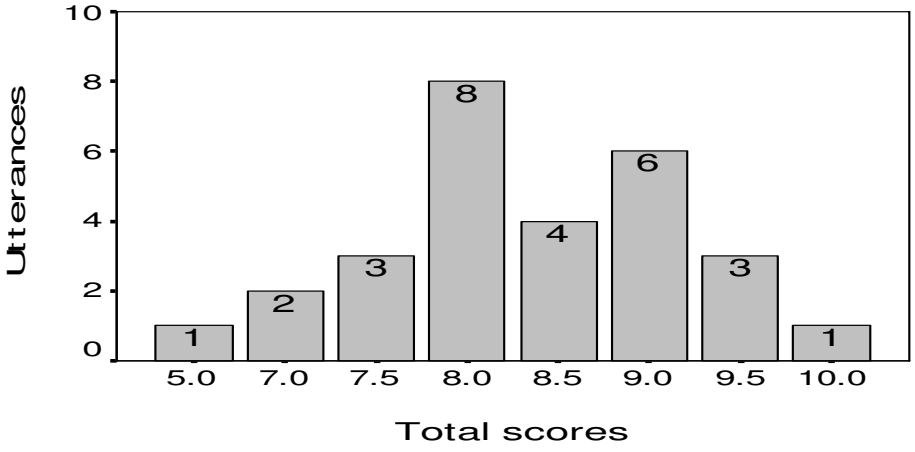

Figure 7 - Total scores for the utterances of critical thinking in the first discussion

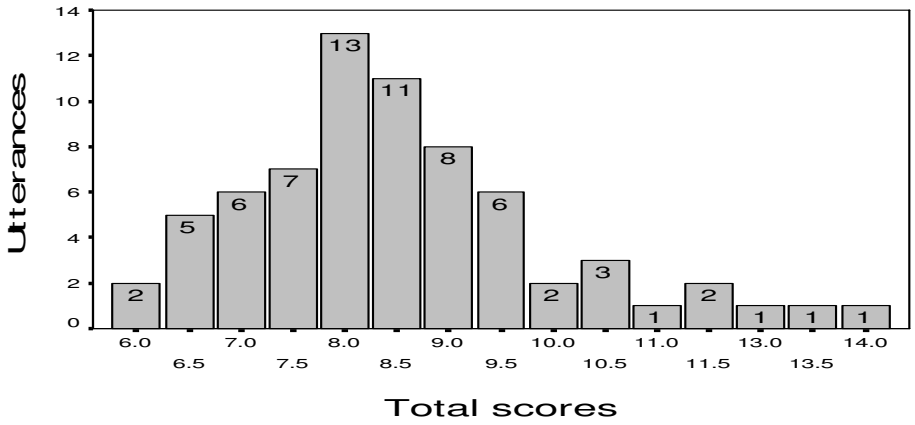

Figure 8 - Total scores for the utterances of critical thinking in the second discussion

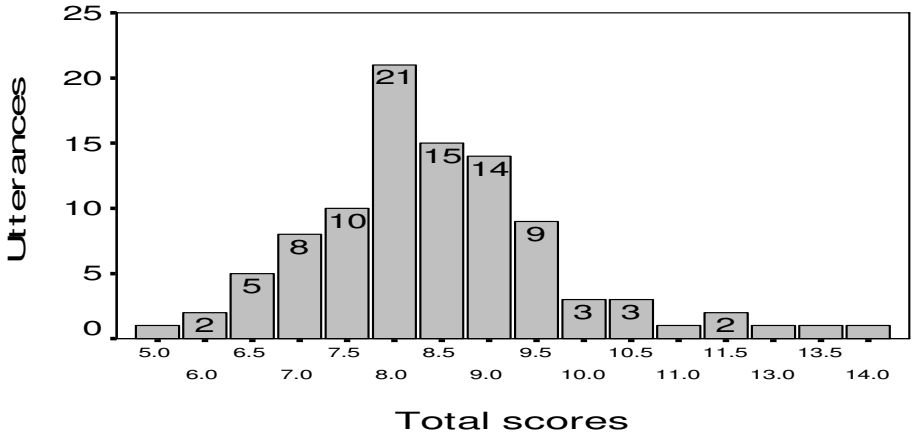

Figure 9 - Total scores for both groups for the utterances of critical thinking

In the first classroom, which had fewer participants and lasted for a shorter time period, it is interesting to note that the grades are spread over a very narrow range, between 5 and 10, the most frequent grade being 8. In the second classroom, the range is wider, between 5 and 14. Still, the most frequent grade is 8. Not surprisingly, in total, the most frequent score for the nature of language is 8 again.

Correlations of the scores

While analyzing the scores given for the different aspects of the criteria used to examine the nature of language used in expression of CT, the researcher was curious about the relationship between the effectiveness scores and the two overtly language-related aspects of the criteria, and therefore decided to look at the correlations of the scores. Table 3 below shows the correlations among all the aspects of the criterion.

		vocabulary score	grammar score	effectiveness in conveying the message score	total nature of language score
vocabulary score	Pearson Correlation	1	.706(**)	.322(**)	.802(**)
	Sig. (2-tailed)	.	.000	.001	.000
	N	97	97	97	97
grammar score	Pearson Correlation	.706(**)	1	.369(**)	.842(**)
	Sig. (2-tailed)	.000	.	.000	.000
	N	97	97	97	97
effectiveness in conveying the message score	Pearson Correlation	.322(**)	.369(**)	1	.755(**)
	Sig. (2-tailed)	.001	.000	.	.000
	N	97	97	97	97
total nature of language score	Pearson Correlation	.802(**)	.842(**)	.755(**)	1
	Sig. (2-tailed)	.000	.000	.000	.
	N	97	97	97	97

Table 3 - Correlations among all the aspects of the criterion.

The table shows that the vocabulary score correlates highly with the grammar score (.706, $p<0.1$) and the total scores (.802, $p<0.1$). However, it seems that the correlation between vocabulary and effectiveness in conveying the message scores is positive, but not as high (.322, $p<0.1$). Therefore, it seems that some other factors, besides vocabulary, are influencing the effectiveness score. It can also be seen from the table that just like vocabulary, grammar also correlates highly with vocabulary

(.706, $p<0.1$), and the total score (.842, $p<0.1$), but its correlation with effectiveness is not very high (.369, $p<0.1$). Again, this means that something else besides grammar is influencing the effectiveness score. This may lead to the conclusion that whether the message is conveyed effectively or not may not depend on effective grammar and vocabulary. The external factor that affects the understanding of the message could be anything, but, according to the researcher, the most likely one is whether there are enough contextual clues in the utterance in question or the ones around it to convey the message or not. The last point is that, effectiveness correlates with total score reasonably well (.755, $p<0.1$).

Other features of the CT utterances

As the researcher read and re-read the discussions, it was observed that in the utterances that involve CT, discourse markers (Stapleton, 2002) seemed to appear much more often than in the other parts of the discussions. In order to determine whether this pattern held true throughout the discussions, all the utterances were examined for such discourse markers, starting with a list of discourse markers identified in Stapleton, and expanding the list as necessary.

These discourse markers include:

I think, for example, like, what about…?, because (of), the reason can/may be, that's why, since, for this, the reason for this, the main reason (is), as you see, in my opinion,	so, this means, I mean, I don't mean, to start with, firstly, secondly, first of all, also, then, another (problem) is that, I don't agree, I agree,	I claim, I recommend, I advise, I am (we are) against…, some people think/claim that..., as I said before, the other thing about…, I will say two main (things), in contrast, but, the point is, actually in fact

When the utterances labeled as CT were examined for these discourse markers, it was revealed that an average of two target discourse markers occurred per CT utterance. In contrast, there are only 1.5 target discourse markers per unclear utterance, 1.15 target discourse markers per undecided utterance and 0.39 target discourse markers per non-CT utterance. As can be seen, CT utterances appear to have many more of these types of discourse markers than the other categories of utterances. Some of these discourse markers will be exemplified in context in the extracts below.

Extract 20

> In my opinion, it start with, it start with the ... this developing technology. Because while thinking situation, before the television comes there were no celebrities, in ... eee … there were no paparazzis. After ... the television was created at the start of this century. Some of people think that they use a program which…eee ... firstly, in fact … we start television every person, in public, so the communication in the neighbourhood and the communication

in friendship is ... eee … decreased. So, the only social … the only free time ... in social ... level is only the..only sit … sit in front of television and see ... look at the television, what they want, what they look. For example, in first the television come to our country there are only 2-3 hours with ...shows something and everyone wait that, but today ... twenty-four hour a day it's living ... and television … paparazzis sell a data from the celebrities privacy. And we only look at eee ... them, because we have to … eee … we have a limited area to spend our time…with social life.

Extract 21

I think we prevent … eee ... we want to prevent invasion of privacy. First of all change the people's minds ... Because if ... if ... if we change the people's minds … they are no interested in the life of the celebrities. So the invasion of the privacy is cancelled.

Extract 22

But I don't think that paparazzis are doing their job so well, because they have already passed the line they should never come close. And they really become an … they really started to annoy all the celebrities. And even all .... eee .... even most of the public. So I ... eee … eee ... recommend them to stop what they have to stop, stop the line and don't pass the line. If you are aware of your responsibilities ... (laughter) … everybody will be happy in the public.

Extract 23

Eee I think ee and eee another problem is that, lots of people ... eee millions of problems in Turkish public doesn't have a job, so their minds are washed with these issues and they are isolated from our ... maybe government problems ... or eee ... country problems, So I think we should make public aware of this and we should work with pub. Paparazzis celebrities … should make a platform and we should eee ... make comments on this

Extract 24

Yea but, what is the sense in that? Like it's not their real personality … like ... nobody cares about their real personality. I mean they just want the life which you shoot.

Extract 25

But the point is … eee ... if ... eee ... some people wants to ... eee .... wants their privacy to be invaded that won't mean an invasion. And there is no problem. Everyone is happy. You earn money, we earn money and there won't be a problem. But if we are annoyed of the situation and ...we want you to stop, you should stop. And ethics require you to stop

In the extracts above, the frequency of discourse markers can be seen clearly. The discourse markers seem to be used for signposting and connecting the ideas together. The students seem to be trying to make their organization more obvious to the listeners with the help of the discourse markers.

Other features of non-CT utterances

Like the discourse markers, another instance of data that emerged as the researcher analyzed the discourse is the tonal and structural difference between the CT and non-CT discourse. Of the 86 utterances of non-CT in total, 32 have an assertive tone, which was manifested in very strong words or phrases. This is, in a way, not surprising, as in the non-CT part of the framework that the researcher designed and used, there are the categories of *accusational or defensive statements* and *emotional statements* which are usually expressed in an assertive tone. Some examples of this assertive tone will be shown in extracts below.

Extract 26:

> Yes but they shouldn't also blame us… They try to… celebrities and public try to put the blame on us.

Extract 27:

> I'm fed up with this...

Extract 28:

> If you don't want, don't watch!

Extract 29:

> Why are you afraid?

Extract 30:

> Because they don't, they can't show their own feelings because of you. And because you change their feelings and you give us what you want to say and

> also. You passed your lines and this means war and you declared the war. Innocent people will harm. So we are defend we are trying to defend our children from your bad effects and also from celebrities bad effects, I have to say this, but you declared the war.

Extracts 26, 27, 28, 29 and 30 include typical examples of non-CT utterances, which include an assertive and/or accusational tone using strong and striking words and expressions like *blame, fed up, don't, afraid, because of you, declare war, harm, defend* and so on. Extracts 5 and 6, which were given previously as examples of non-CT utterances, also include the same kind of characteristic tone, using such phrases as, such as *makes me angry, they started the situation.*

In addition, as can also be seen in the example utterances, non-CT utterances are usually much shorter than CT utterances. The average number of words per non-CT utterance is 10.63, whereas the average number of words per CT utterance is 47.43. The reason for this could be that when the students were expressing critical thinking, they had to support what they said with further examples, evidence and so on, which lengthens the utterances.

In summary, looking at the utterances in terms of three aspects of the nature of language, both individually and combined, reveals that the average scores fall into the middle of the ranges. Therefore, it might be concluded that the students are not very effective in vocalizing their CT in L2. In addition, it was seen that the vocabulary and grammar correlate with each other relatively highly, whereas they do not correlate with the score of effectiveness in conveying the message as highly. Furthermore, it was revealed that the utterances which include CT use discourse markers frequently and correctly, which may lead to a better conveyance of the organization and content of the message. Last, non-CT utterances are much shorter

than CT utterances, as they are generally left unsupported, and they include an assertive tone owing to the strong words used in them.

Conclusion

In this chapter, the results of the analysis of the discourse were presented. The results included the amount of CT expressed in individual discussions and in total, and the nature of language in utterances including CT in terms of vocabulary, grammar, effectiveness in conveying the message and the overall nature of language as well as the correlations among these scores. In addition to that, some incidental results about features of CT and non-CT expressions were analyzed. In chapter five, the major findings of the study, the pedagogical implications and the limitations of this study will be described.

CHAPTER V: CONCLUSION

Introduction

This study was designed to investigate the ability of Turkish EAP students to express CT in their oral L2 discourse in discussion tasks. It attempted to address the following research questions:

1. What is the amount of critical thinking expressed in class tasks that invite critical thinking in EAP classrooms in Turkish universities?
2. What is the nature of language when vocalizing critical thinking in class tasks that invite critical thinking in EAP classrooms in Turkish universities?

This study was conducted with two classes of EAP students studying at Middle East Technical University. Two classes were video-recorded while engaged in a task that required discussion. Three groups of students discussed the issue of invasion of privacy, representing paparazzi, the public and celebrities. The first discussion was about 35 minutes long and had 10 participants. The second discussion was about 60 minutes long and had 24 participants. Then, the recordings were transcribed and analyzed for the amount of CT expressed and the nature of the language used in the utterances that included CT.

This chapter includes a discussion of the research findings concerning the research questions and the relevant literature, the limitations of the study, the pedagogical implications derived from the results and suggestions for further study.

General Results and Discussion

The major findings of this study will be presented in two sections: the amount of CT expressed and the nature of oral language in tasks that invite CT.

What is the amount of CT expressed in class discussions?

One major result that emerged from initial attempts to analyze the discourse of these discussions was that spoken discourse, especially the discourse of impromptu talk, is much different from the discourse of online discussions or written discourse in terms of the expression of CT. Although it is possible to observe the same categories that have been described in other studies of critical thinking in oral discourse, it may not be possible to identify them as different chunks. This difficulty was experienced by the researcher when attempting to categorize the instances of critical thinking with a second rater. Online discussions, where moderators manage the discussions, learners usually discuss around predetermined questions, and/or they have time to think about what they are going to say (see for example Uzuner, 2007), appear to have a neater and more organized discourse. The utterances appear to be much more disorganized in oral discourse, usually seeming to include more than one category of CT, which are not separable from one another or are open to interpretation. Not surprisingly, the oral expression of CT in these classroom discussions was also different from problem-solving discourse, in which there are certain stages that students are expected to follow, and therefore a much neater discourse is obtained (Kamin, O'Sullivan, Younger, & Deterding, 2001). The utterances on which an agreement could not be reached by the researcher and the second rater in the discussions may also serve to illustrate the chaotic and complicated nature of spoken L2 discourse. Therefore, it can be suggested to even

have a third rater for such utterances for further studies in the same field. A third rater might resolve the disagreements between the two raters and in this way, the number of utterances on which an agreement could not be reached can be decreased, and thus, the reliability rate can be increased.

Analysis of the discussions indicated that, in both classrooms together, 45% percent of the utterances were identified as containing critical thinking whereas 40% of them were identified as non-CT. These percentages are quite similar to each other. However, a very important point to note here is that it is not possible to determine whether this amount is expected or not, since there is not any stated amount of CT expression in the literature to set the standard in oral language. Meyer (2003) found that 80% of online discussions consisted of higher order thinking skills, but a search of the literature revealed no study that provided any measure of critical thinking in spoken language. Still, it can be said that it may not be possible or even desirable for every utterance to consist of CT. It can surely be said that some percentage of a discussion has to consist of other kinds of conversational moves or functions, such as "rapport-building" in Kamin et al.'s (2001, p. 30) category of *group process issues* or "inviting others to talk" as in Uzuner's (2007, p. 405) *invitational* category in her educationally less valuable talk framework. However, in the absence of any previous study of the amount of critical thinking expressed in L2 classroom oral discourse, this study might be said to have set a benchmark to be referred to by other studies in the future.

The overall picture of the amount of CT expressed in the two classrooms shows that the amount of CT expression differs from classroom to classroom in spite of the fact that the task is the same and the classrooms are in the same university, and

thus, in the same learning environment and culture. The figures show that in the first classroom much less CT (34% versus 53%) and much more non-CT (54% versus 31%) was expressed than in the second classroom. The difference in the amount of expression of CT may be a result of the individual backgrounds of the students that constitute these classes. Individual factors might include the high schools they attended, their majors, self-confidence, dispositions to critical thinking and even their level of English.

That much less CT was expressed in the first classroom might also be due to several other factors. First, the groups' sizes might have had an effect. There were only ten students in the first classroom, which resulted in three students in a group, with one student being the jury, whereas there were 24 students in the second classroom, which allowed six to seven students in a group, with four students being the jury members. The result that the first classroom had less expression of CT than the second classroom supports Rau and Heyl's conclusions (1990, as cited in Gokhale, 1995, p. 25) that "smaller groups (of three) contain less diversity; and may lack divergent thinking styles and varied expertise that help to animate collective decision making". Thus, it can be said that small groups may not be as inspiring in terms of CT as bigger groups. Second, the students in the first classroom did not seem as interested and motivated as those in the second group. Therefore, in the preparation time given for the discussion, it appeared that they did not talk and share ideas as effectively as the students in the second classroom. These results are compatible with a highly accepted theory in the literature that CT benefits from interaction, collaboration and "active exchange of ideas" (Gokhale, 1995, p. 22). Third, the number of students in the classes might also have had an effect on the

amount of CT expressed. From a common sense point of view, the higher the number of students, the higher the chances of having individuals who are able to express their critical thinking in the group. From the same perspective, the longer the discussion, the higher the chances may be of critical thinking occurring. Thus, the fact that it lasted almost twice as long as the first discussion may explain why more CT was expressed in the second discussion. Last, two different teachers managed the discussions differently, which might have had an effect on the way the discussions unfolded. The teacher of the second classroom was less intrusive, almost never interrupting the students. Rather, she just nominated those that raised their hands and recapped the points made only after utterances that were not understood by the class well, in order to keep the flow of the discussion. In this way, the students may have been able to find more room than those in the first discussion to express and support their opinions with further evidence, examples and arguments, without the fear of time constraints, and her behavior might have had an effect on the amount of CT expressed in the students' discourse. In contrast, the teacher in the first classroom was much more dominant in the classroom and she made it obvious that she was the leader of the discussion. Also, she frequently drew her own conclusions from what students said, which I believe interfered with the natural flow of the discussion. After all, it is known from the literature that the behavior of the teacher influences the nature of the talk in the classroom. For example, Tobin (1987) found that teacher's wait time after asking a question had an impact on the cognitive complexity of the discourse produced by the student.

What is the nature of language when vocalizing CT in class tasks that invite CT?

The data revealed that ten utterances received the lowest vocabulary score of 2, while 13 utterances ranked at 2 and below in the scores of effectiveness in conveying the message. However, 19 utterances received grammar scores of 2 and below. Furthermore, the most frequent grade given for grammar was 2.5, which is lower than the most frequent grades given for the other components of the criterion, and the mean for grammar scores is lower than the means of the other aspects in the criterion (2.716 vs. 2.845 and 2.928) (see table 2 on page 49). From this distribution of the scores, it appears that the students have more trouble with grammar than with the other aspects of the criterion in expressing CT. Vocabulary seems to be less of a problem than grammar. That students were able to express their ideas and understand each other in spite of their problematic grammar is in line with what Wilkins (1972, p. 111) states: "while without grammar very little can be conveyed, without vocabulary nothing can be conveyed". It can be assumed that they may not have been able to express that "very little" if vocabulary was their most problematic area.

On the other hand, considering the most frequent grades given to each aspect of the criteria, grammar may not be their only problematic area. The most frequent grade given in vocabulary is 3, in grammar it is 2.5, and in effectiveness in conveying the message it is 3 out of 5. This results in a total score of 8 out of 15. These numbers are very similar to each other in that they are either just half of the total grade or slightly above it. This may indicate that the students are average in their abilities to use vocabulary and grammar and average in their effectiveness in conveying their message, and as a result, they achieved only average total scores when externalizing critical thoughts. This may also lead us to the conclusion that the

students are not at the level that they are assumed by the school in oral language, because, as American Council on the Teaching of Foreign Languages (ACTFL) defines, advanced level students can:

> perform all Advanced-level tasks with linguistic ease, confidence and competence. They are able to consistently explain in detail and narrate fully and accurately in all time frames... They can provide a structured argument to support their opinions, and they may construct hypotheses, but patterns of error appear. They can discuss some topics abstractly, especially those relating to their particular interests and special fields of expertise, but in general, they are more comfortable discussing a variety of topics concretely. Advanced-High speakers may demonstrate a well-developed ability to compensate for an imperfect grasp of some forms or for limitations in vocabulary by the confident use of communicative strategies, such as paraphrasing, circumlocution, and illustration. They use precise vocabulary and intonation to express meaning and often show great fluency and ease of speech. (1999, p. 3)

This means that the students, who are considered to be advanced students of English in English 102, should have had higher than average language skills in all aspects of the nature of language criteria. This could be due to several factors including the fact that the EAP courses in which the study was conducted do not test speaking yet, the speaking objectives of the university for impromptu talk are not very ambitious, (*Middle East Technical University School of Foreign Languages Curriculum Policy Document/Interim Report*, May, 2004) and the traditional way of lecturing in Turkish universities, especially in technical universities, that does not require the students to speak much, and thus does not allow them to improve their speaking skills as much as their reading, listening and writing skills

In addition, two figures in this picture may lead us to two complementary results. First, the percentages of the most frequent grades, which are just half or just above half of the total grade, in grammar (38%) and vocabulary (41%) are much higher than the percentage of the equivalent most frequent grade in effectiveness in

conveying the message (28%). Second, the percentage of the grades that are higher than 3 in effectiveness in conveying the message (32%) is considerably higher than the percentage of those in vocabulary (16%) and grammar (9%) scores. In addition to that, the mean score for effectiveness in conveying the message (2.928) is higher than the means of vocabulary (2.845) and grammar (2.716) scores. From these results, one may conclude that the students are less successful in vocabulary and grammar than they are in overall effectiveness in conveying the message, which, in turn, means that they can convey their message in spite of their lack of sufficient grammar and vocabulary. This indicates that they might be achieving some of what is described in the ACTFL level description, in terms of "paraphrasing, circumlocution and illustration" (1999, p. 3) to compensate for their insufficient grammar and vocabulary. Furthermore, the fact that the grammar and vocabulary scores did not correlate as highly with the effectiveness scores also indicates that effectiveness is not necessarily a matter of accuracy.

The overall picture of the nature of the language used while expressing CT, including both classrooms, is not very encouraging. The scores at the highest end of the scale (10-15) constitute only 12% of the total grades given. These figures indicate a real need for more effective oral language for the students to express their critical thinking in L2.

It is also important to notice that the scores of the individual classrooms suggest that the two classrooms do not differ from each other vastly in the nature of the language used. The most frequent score in both classrooms is 8, which is just above half of 15, and they both have 5 as their lowest score. The wider range of grades in the second discussion might be associated with the fact that there were

more students in that classroom. Also, although the second discussion had high scores, such as 11, 11.5, 13, 13.5, 14, in contrast to the first discussion, there were only six utterances that were given these grades in total, belonging to four students, who might be considered to be exceptional cases as they might have been abroad and might be able to speak remarkably better than the average student.

The general analysis of the discourse features of the discussions indicates that the students can use frozen expressions in debate language, such as *firstly, secondly, however, in my opinion, I agree, I am against* and the like, appropriately and correctly. Although their study was not about student talk, Eslami, Eslami-Rasekh and Data (2007) found that discourse markers contribute to academic lectures' comprehensibility to a great extent, for example. Therefore, their use could be an explanation of how the students can express their message better than they might be expected to with the insufficient vocabulary and grammar they have. They signpost well, and thus, they give the listeners a lot of clues to guess the meaning from the context. However, as their scores for effectiveness in conveying the message are only average, students could use them more effectively. Therefore, teaching discourse markers should not be underestimated in EAP classes. Also, as it was seen that using discourse markers effectively does not necessarily mean that they can explain the content as well, it is possible to say that there is definitely more to expressing CT than being able to use discourse markers well in speech. After all, although not as often as in CT utterances, the discourse markers were used in the non-CT and unclear utterances as well. As indicated before, 1.5 target discourse markers were used per unclear utterance and 0.39 target discourse markers were used per non-CT utterance.

Yet another result of the analysis of the discourse of the discussions is that it appears that non-CT discourse usually involves short sentences or questions with a striking or assertive tone to convey the message. In contrast, CT discourse seems to involve much longer statements, as the speaker has to support his/her argument in a satisfactory manner. From this point, one may conclude that these characteristics can be tentatively considered markers of CT expressed in oral discussions in a second language and may contribute to the further development of a framework of critical thinking in such contexts.

Limitations

This research had to be completed in a very limited amount of time. Therefore, it had to be conducted within the constraints of a small-scale study. In principle, if there had been more classes involved in the research process to increase the amount of data obtained, the results could be more generalizable. Also, the fact that the researcher divided the discussions into utterances herself, without a second rater, also due to time constraints, might also be considered to be a limitation.

Other factors that might have affected the results include, first, the different class sizes and the different lengths of the discussions. A second factor could be the presence of the cameras, which might have caused the students to talk more than they usually do or, conversely, intimidated them. The fact that the utterances could not be categorized further than CT or non-CT can also be considered to be a limitation, as well as the number of the utterances on which the raters could not agree.

Implications

The results of the analysis of the amount of CT expressions in the discussions show that 45% of the utterances in the discussions included CT. This indicates that students can express their critical thinking skills in English, although they and their teachers frequently claim the opposite. However, the fact that the non-CT utterances also constitute a significant percentage of the discussions (40%) may be causing teachers to notice only language that does not represent critical thinking. Also, the fact that the students achieved only average scores in terms of vocabulary, grammar and effectiveness in conveying the message may be causing misunderstandings of the ideas by the teachers and may justify students' not feeling comfortable in situations where they need to engage in CT and express it in English. EAP courses alone cannot increase the amount of CT expressed. However, what EAP should do is to give students more effective language to enable them to express their CT more comfortably.

One suggestion for this purpose could be to have students get involved in projects long and complex enough to practice higher-level reasoning in L2 so that "the rhetorical and argumentation conventions can be identified, studied, practiced, and questioned" (Pally, 2001, p. 299). In other words, students might benefit from sustained content-based EAP courses which deal with one or, at most, two topics for the whole semester. Another suggestion would be to focus on vocabulary learning strategies in classes more explicitly and more frequently in the classes. This might help students learn the less commonly used abstract and academic vocabulary items that they need to explain themselves in more comprehensible and precise ways. Incorporating more practice requiring spontaneous oral externalization of CT skills,

such as debates and discussions, might also help to give students more opportunities to express their opinions on the topics of English classes orally and might have them develop the habit of speaking up. Group learning activities could be used as the main educational tool to increase the effectiveness of these discussions, and students could be taught how to talk in a group in an "exploratory" manner, in Swan's terms (Swan, 2007, p. 342), to increase knowledge building discourse (Suthers, 2001), instead of participating in a discussion in an egocentric way to defend their own opinions at the expense of others, which often may lead to a non-critical way of thinking and talking.

Students should also be given more opportunities to practice accurate grammar in EAP classes. Perhaps, as well as content, the medium, which is language, should be focused on in the class. The current practice in freshman EAP courses, of not addressing grammar directly and leaving the students to improve their grammar on their own, is not satisfactory. Perhaps, the students need to continue to focus on grammar, even in these courses.

Another educational implication of this study relates to the sizes of the groups and the classroom. It appears that the second classroom, which had 24 students, with six to seven members in each group, was more motivated and interested in the discussion and therefore their oral expression of CT was greater than the students in the first classroom, which had 10 students, with only three members in each group. It might work to keep the group sizes a little bigger than normal in activities that require CT to ensure diversity of opinions and critical thinking skills (Rau & Heyl, 1990, as cited in Gokhale, 1995).

In addition, discourse markers might be helping the students as they are frozen chunks of L2 in their minds. They are ready to be used in a spontaneous talk situation, which requires a lot of complex cognitive processes and thus use of these frozen chunks might be saving students from some of these complex processes (Sinclair, 1991). In a task that invites CT, these expressions might even be more helpful as students are dealing with complex thought in addition to the complexities of the situation. Therefore, teachers can spend regular time in the class to teach the discourse markers of all kinds in their classrooms as this is very likely to be a feature of CT discourse in oral language in L2. Other types of frozen language such as collocations might also be considered in the same respect.

Yet another implication is the topic choice. As an explanation for why students in the first classroom were not as interested as the students in the second classroom, it might be said that the first group might not have found the topic as interesting as the second group. Their expression of CT might have been lower than the students in the second classroom because of their lack of motivation about the topic. Thus, it can be said that it is important to find the topics that the target student profile would be interested in to increase CT expression

The teacher's style of leading the discussion might also have had an effect on students' expression of CT. The existence of the teacher as a strong authority figure and the leader of the discussion might hinder the expression of CT. As Neilsen advises (1989, as cited in Tama, 1989) teachers should "allow learners to be actively involved in the learning process" and "establish a supportive learning environment that respects student opinions while giving enough direction to ensure their relevance to a topic, and to provide ample opportunities for learners to collaborate" (p. 4).

Therefore, students should be leaders and the followers in the discussion. They should be responsible for their answers and ideas. They should also be responsible for making sure that they are not repeating themselves as this is generally the reason why teachers tend to interrupt. They should not be given the comfort of having an authority figure in the discussion, who they know will intervene whenever there is a problem, fix it, put things in order, and increase the level of the discussion for them. In brief, it is possible to advise teachers to let students accept the full responsibility of their own discussion and ideas they put forward.

Suggestions for further research

Based on the limitations of the current study, a further study could be conducted with more than two classes to see if the discourse of CT in L2 shows the same patterns in a wider field. Similarly, it can be conducted with more similar classes in terms of class and group sizes. Also, a second rater could be used in identifying the utterances to increase the reliability.

There could be an attempt to identify the categories of critical thinking with a more general framework such as Stapleton's (2002, pp. 537,538), which includes argument, evidence, recognition of opposition and refutation and fallacies for the writing contexts. Such a framework could make spoken discourse more categorizable in terms of CT, after adaptation to speaking contexts.

Conducting the same study with the same students both in their native-languages and in English would also be interesting to see the similarity or the difference between the amounts of expression of CT. Such a study might help understand if the language affects the expression of CT or not. Also, the study could

be carried out in different cultures to see the cross-cultural differences or similarities in CT expression.

In addition, different classroom tasks could be evaluated with the same kind of a study to investigate and compare their abilities to invite CT expression in an EAP classroom. In this way, the kinds of tasks that enhance CT expression in a foreign language could be determined and used in EAP classes.

Conclusion

The primary aim of this study was to find out the amount of CT expressed in tasks that invite CT in EAP classes in EFL settings and to find out the nature of students' language while doing so. To this end, two classrooms were video-recorded while they were doing a task that invites CT. Then, the discussions were transcribed and the discourse in them was analyzed. What was learned from this study is that the students have the potential to express their CT in their L2 and, indeed, they are doing so to a certain extent, in contrast to what the teachers claim. However, they have difficulty in doing so as a result of their lack of necessary language, just as they claim. The students' critical thinking is discernible in their speech although it may not be very obvious. The findings of this research revealed the need for further support for students to vocalize their CT skills more effectively in spoken L2. This support might include putting more emphasis on vocabulary learning strategies, having more practice on debates and discussions in class, putting more emphasis on accuracy in the classes and paying attention to group sizes, topic choice, and discussion leading style when employing tasks that invite CT in the classroom.

REFERENCES

Alagozlu, N. (2007). Critical thinking and voice in EFL writing. *Asian EFL Journal, 9*(3), 118-136.

Anderson, T., Howe, C., Soden, R., Halliday, J., & Low, J. (2001). Peer interaction and the learning of critical thinking skills in further education students. *Instructional Science, 29*(1), 1-32.

Atkinson, D. (1997). A critical approach to critical thinking. *TESOL Quarterly, 31*(1), 71-94.

Atkinson, D., & Ramanathan, V. (1995). Cultures of writing: An ethnographic comparison of L1 and L2 university writing/language programs. *TESOL Quarterly, 29*(3), 539-568.

Bailin, S. (1998, August). *Skills, generalizability and critical thinking.* Paper presented at the Twentieth World Congress of Philosophy, Boston, Massachusetts.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Common misconceptions of critical thinking. *Journal of Curriculum Studies, 31*(3), 269-283.

Bailin, S., & Siegel, H. (2003). Critical thinking. In N. Blake (Ed.), *Philosophy of education* (pp. 181-193). London: Victoria Blackwell Publishing.

Barnet, S., & Bedau, H. (2002a). *Critical thinking, reading, and writing: A brief guide to argument*. Boston: Bedford/St. Martin's.

Barnet, S., & Bedau, H. (2002b). *Current issues and enduring questions: A guide to critical thinking and argument, with readings.* Boston: Bedford/St. Martin's.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals.* New York: David McCay.

Bloom, B. S., Thomas, J. H., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw Hill.

Bridges, D. (1993). Transferable skills: A philosophical perspective. *Studies in Higher Education, 18*(1), 43-51.

Brookfield, S. D. (1987). *Developing critical thinkers*. San Francisco: Jossey-Bass.

Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.

Brown, H. D. (1972). Cognitive pruning and second language acquisition. *The Modern Language Journal, 56*(4), 218-222.

Carter, R., & Nunan, D. (2002). *The Cambridge guide to teaching English to speakers of other languages*. Cambridge: Cambridge University Press.

Cuban, L. (1984). Policy and research dilemmas in the teaching of reasoning: Unplanned designs. *Review of Educational Research 54*(4), 655-681.

Day, R. R. (2005). Review: Critical Reading in Language Education. *English Language Teaching Journal, 59*(3), 258-259.

Dewey, J. (1910). *How we think*. Boston: Heath & Co. .

Dewey, J. (1938). *Experience and education*. New York: Touchstone.

EAP 1121-1620 reading competencies, levels 1-6 (2000). Retrieved 5 August 2008 from www.mdc.edu

EAP at the University of Prince Edward Island. (2003-2004, 29 May 2006). Retrieved 5 August 2008, from http://upei.ca

Elsegood, S. Teaching critical thinking in an English for academic purposes program using a "claims and supports" approach, Refereed Paper. Retrieved 1 July 2008, from www.fyhe.qut.edu.au/past_papers/papers07/final_papers/pdfs/4e.pdf -

Ennis. (1981). Eight fallacies in Bloom's taxonomy. In *Philosophy of Education 1980* (pp. 269-273). Bloomington, IL: Philosophy of Education Society.

Ennis. (1993). Critical thinking assessment. *Theory into Practice, 32*(3), 179-186.

Eslami, Z. R., Eslami-Rasekh, A., & Data, B. (2007). Discourse markers in academic lectures. *The Asian EFL Journal Quarterly, 9*(1), 22-38.

Evers, A. (2007). *Does discipline matter? Pedagogical approaches to critical thinking in English for academic purposes and economics.* Unpublished Master's Thesis, Carleton University Ottawa.

Facione, P. A., Sanchez, C. A., Facione, N. C., & Gainen, J. (1995). The disposition toward critical thinking. *Journal of General Education, 44*(1), 1-25.

Fisher, A. (2001). *Critical thinking: An introduction*. Cambridge: Cambridge University Press.

Fisher, A., & Scriven, M. (1997). *Critical thinking: Its definition and assessment*. Norwich, UK: Centre for Research in Critical Thinking.

Flemming, L. (2000). *Reading for thinking* (Third ed.). New York: Houghton Mifflin Company.

Gokhale, A. A. (1995). Collaborative learning enhances critical thinking. *Journal of Technology Education, 7*(1), 22-30.

Kamin, C., O'Sullivan, P., Younger, M., & Deterding, R. (2001). Measuring critical thinking in problem-based learning discourse. *Teaching and Learning in Medicine, 13*(1), 27-35.

Kern, R. G. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and characteristics of language production. *The Modern Language Journal, 79*(4), 457-476.

Klein, C. E. (1993). More than a required skill in today's curriculum: Critical thinking and collaborative learning in foreign languages. *Mid-Atlantic Journal of Foreign Language Pedagogy, 1*, 91-96.

Lee, J., Dineen, F., & McKendree, J. (1998). Supporting student discussions: It isn't just talk. *Education and Information Technologies, 3*(3), 217-229.

Lewis, A., & Smith, D. (1993). Defining higher order thinking *Theory into Practice, 32*(3), 131-137.

Mackey, J. (1977). Three problem-solving models for the elementary classroom. *Social Education, 41*(5), 408-410.

McCarthy, M., & Carter, R. (1994). *Language as discourse: Perspectives for language teaching*. London: Longman.

McLean, C. L. (2005). Evaluating critical thinking skills: Two conceptualizations. *Journal of Distance Education, 20*(2), 1-20.

Meyer, K. A. (2003). Face-to-face versus threaded discussions: The role of time and higher-order thinking. *Journal of Asynchronous Learning Networks, 7*(3), 55-65.

Middle East Technical University School of Foreign Languages Curriculum Policy Document/Interim Report. (May, 2004). Ankara: Middle East Technical University.

Middle East Technical University School of Foreign Languages Curriculum Renewal Project. (2004-2005). Ankara: Middle East Technical University.

Neilsen, A. R. (1989). *Critical thinking and reading: Empowering learners to think and act*. Bloomington, Indiana: ERIC Clearinghouse on Reading and Communication Skills.

Newman, D. R., & Johnson, C. (1997). Evaluating the quality of learning in computer supported co-operative learning. *Journal of the American Society for Information Science, 48*(6), 484-495.

Newman, D. R., Webb, B., & Cochrane, C. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology, 3*(2), 56-77.

Norris, S. P. (1985). Synthesis of research on critical thinking. *Educational Leadership, 42*(8), 40-45.

Pally, M. (2001). Skills development in 'sustained' content-based curricula: Case studies in analytical/critical thinking and academic writing. *Language and Education, 15*(4), 279-305.

Paul, R. (1990). Critical thinking: What, why, and how. In *Critical thinking: What every person needs to survive in a rapidly changing world*. Sonoma: Center for Critical Thinking and Moral Critique Sonoma State University.

Paul, R. (2007). *Critical thinking in every domain of knowledge and belief.* Paper presented at the 27th Annual International Conference on Critical Thinking, Berkeley, CA.

Paul, R., Binker, A. J. A., Martin, D., Vetrano, C., & Kreklau, H. (1989). *Critical thinking handbook: 6th- 9th grades*. Rohnert Park, CA: Center for Critical and Moral Critique, Sonoma State University.

Paul, R., Elder, L., & Bartell, T. (1997). *California teacher preparation for instruction in critical thinking: Research findings and policy recommendations*. Sacramento, California: Foundation For Critical Thinking.

Paul, R., Fisher, A., & Nosich, G. (1993). *Workshop on critical thinking strategies*. CA: Foundation for Critical Thinking, Sonoma State University.

Rasool, J., Banks, C., & McCarthy, M. (1993). *Critical thinking, reading and writing in a diverse world*. Belmont, California: Wadsworth Publishing Company.

Risinger, C. F. (1987). Improving writing skills through social studies. ERIC Digest No. 40. Bloomington, Indiana: ERIC Clearinghouse for Social Studies/Social Science Education.

Rubin, J. (1979). What the "good language learner" can teach us. In J. B. Pride (Ed.), *Sociolinguistic aspects of language learning and teaching* (pp. 17-26). London: Oxford University Press.

Ruland-Parker, J. R. (1999). *Relationship of classroom environment to growth in critical thinking ability of first year college students.* State University of New York, New York.

Schamber, J. F., & Mahoney, S. L. (2006). Assessing and improving the quality of group critical thinking exhibited in the final projects of collaborative learning groups. *JGE: The Journal of General Education, 55*(2), 103-137.

Shermis, S. S. (1999). *Reflective thought, critical thinking (Digest #143).* Bloomington, Indiana: ERIC Clearinghouse on Reading, English, and Communication

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Smith, C. B. (1991). *A commitment to critical thinking*. Bloomington: Grayson Bernard Publishers.

Stahl, N. A. (1991). *How college learning specialists can help college students*. Bloomington, Indiana: Eric Clearinghouse on Reading and Communication Skills.

Stapleton, P. (2002). Critical thinking in Japanese L2 writing: Rethinking tired constructs. *ELT Journal, 56*(3), 250-257.

Suthers, D. D. (2001). *Collaborative representations: Supporting face to face and online knowledge-building discourse.* Paper presented at the 34th Hawaii International Conference on System Sciences.

Swan, J. (2007). Designing educationally effective discussion. *Language and Education, 21*(4), 342-359.

Tama, M. C. (1989). *Critical thinking: Promoting it in the classroom*. Bloomington, Indiana: ERIC Clearinghouse on Reading and Communication Skills.

Tobin, K. (1987). The role of wait time in higher cognitive level learning. *Review of Educational Research, 57*(1), 69-95.

Uzuner, S. (2007). Educationally valuable talk: A new concept for determining the quality of online conversations. *MERLOT Journal of Online Learning and Teaching, 3*(4), 400-410.

Vermillion, D. M. (1997). *Developing critical thinking skills in EAP students.* Unpublished Master's Thesis, Biola University.

Wallace, C. (2003). *Critical reading in language education*. Hampshire: Macmillan.

Wilkins, D. A. (1972). *Linguistics in language teaching*. London: Edward Arnold.

Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research, 72*(2), 131-175.

APPENDIX A: CRITERIA FOR NATURE OF LANGUAGE

VOCABULARY	**Very poor** (Very simple and limited vocabulary, frequent wrong use of vocabulary, hinders understanding)	**Poor**	**Average** (Basic vocabulary, used mostly correctly, doesn't hinder understanding)	**Good**	**Excellent** (Rich, diverse, correct, extensive use of vocabulary)
	1	*2*	*3*	*4*	*5*
ABILITY TO USE GRAMMAR CORRECTLY	**Very poor** (Basic grammar , simple sentence structures, mistakes, frequent errors, hinders understanding)	**Poor**	**Average** (Applies rules of grammar but lacks control , occasional advanced grammar, mostly basic grammar, errors do not hinder understanding)	**Good**	**Excellent** (A balance of simple and advanced sentence structures, used correctly, able to explain precise points or subtle differences in terms of ideas with the help of grammar)
	1	*2*	*3*	*4*	*5*
EFFECTIVENESS (ABILITY) IN CONVEYING THE MESSAGE	**Very poor** (Ineffective use of language, very difficult to understand the idea without a lot of interpretation.)	**Poor**	**Average** (Quite effective use of language, pretty understandable.)	**Good**	**Excellent** (Excellent use of language, ideas are very clearly stated with the help of the language.)
	1	*2*	*3*	*4*	*5*

APPENDIX B: MIDDLE EAST TECHNICAL UNIVERSITY HUMAN SUBJECTS ETHICS COMMITTEE APPLICATION FORM

Human Subjects Ethics Committee Application Form

Studies conducted in Middle East Technical University (METU) and/or studies conducted by METU personnel/students, which involve collecting data from human participants, are subject to review by the METU Human Subjects Ethics Committee (HSEC). Applicants should submit this application form to the METU HSEC along with the other required documents (see the Application Check List). Approval of the HSEC is required before the start of data collection from human participants.

1. Title of study__**Critical Thinking in a Second Language**_____
2. Type of study (Check the appropriate box)
   □ Academic Staff Study □ Doctorate Thesis ■ Master Thesis □ Other (specify)

_____

3. Researcher's / Researchers':
   Name – Surname **_Buket Esra Tarakçıoğlu**_____ Department: **Modern Languages**_____
   Phone**_2103925**_____
   Address____**Department of Modern Languages Room: 103 Middle East Technical University**
   e-mail address: **tbuket@metu.edu.tr**_
4. Advisor's (or the Supervising Faculty Member's):
   Name – Surname (If applicable)_JoDee Walters Phone __**2901559**_____
5. Expected time frame of the study/project: Start **_3_/_3_/_2008** End **_21/_3/_2008_**
6. Organizations, institutions in which data collection is planned to be accomplished:

a. _Faculty of Engineering_____ e. _____
b. _Faculty of Social Sciences_____ f. _____
c. _____ g. _____
d. _____ h. _____

7. Whether the project is supported/funded or not: ■ Supported □ Not Supported
   If supported, specify institution: ■ University □ TUBITAK
   □ International (Specify)_____ □ Other (Specify) _____
8. Status of the application: ■ New Application □ Revised Application □ Extension of a Previous Project

If it is an extension of a previous project, does the current study show any differences from the previously approved one?

□ Yes □ No

If yes, please explain:

_____

_____

_____

*Undergraduate students conducting research must have an academic advisor/instructor supervising their research.

9. Does the study require giving partial/incorrect information to the participants or keeping them completely uninformed about the purpose of the study? ■ Yes □ No

   If yes, please explain: **The students will not be explained that the aim of the study is to investigate their critical thinking skills as this might affect their behaviors.**

9. Does the study involve questions/items, procedures or manipulations/applications that jeopardize the physical or mental health of the participants? □ Yes ■ No

   If yes please explain

   :_____

   _____

   _____

   Number of participants: **_40_**

10. Will there be a control group? □ Yes ■ No

11. In the list below, please check the items which best describe the participants of the study.

    ■ University Students

    € Employed Adults

    € Currently Unemployed Adults

    € Preschoolers

    € Elementary School Students

    € High School Students

    € Child Laborers

    € Senior Citizens

    € Mentally Handicapped / Challenged People

    € Physically Handicapped / Challenged People

    € Prisoners

    € Other (Please Specify) _____

12. From the list below, please specify the methodology to be included in the study.

- € Survey
- € Interview
- € Observation
- € Administering a test in a computer environment
- € Video/film recording
- € Voice recording
- € Having participants use alcohol, drugs or any kind of chemicals
- € Exposure to high intensity stimuli (light, sound, etc)
- € Exposure to Radioactive Material
- € Other (Please Specify):_____

APPENDIX C: HUMAN SUBJECTS ETHICS COMMITTEE PROJECT

INFORMATION FORM

1. Write a detailed description of your study including your hypotheses.

This study was designed to investigate the ability of EAP students to engage in critical thinking in their oral L2 discourse in class discussions. It attempted to address the following research questions: What is the nature of the thinking done in group work discussions in EFL settings? a. What is the amount of CT in group work discussions? b. What are the types of CT in a group work discussions?

2. Explain the data collection plan, specifying the methods, scales, tools and techniques to be used. (Please hand in a copy of all types of scales and questionnaires to be used in the study along with this document.)

The students studying in an English medium university will perform a discussion task from coursebooks. Their discussion will be video-recorded and coded according to the type and amount of critical thinking using a framework collated by the researcher from the literature results will be reported in terms of frequency and percentages of type of critical thinking, a the quality of critical thinking expressed in the discussions. (a copy of the framework is attached)

3. Write down the expected results of your study.

Not applicable.

4. Does your study involve items/procedures that may jeopardize the physical and/or psychological well being of the participants or that may be distressing for them? If yes, please explain. Specify the precautions that will be taken to eliminate or minimize the effects of these items/procedures.

No.

5. Will the participants be kept totally or partially uninformed of the aim of the study? If yes, explain why. Indicate how this will be explained to the participants at the end of the data collection in debriefing the participants.

The students will not be informed that the researcher is investigating their critical thinking skills. At the end of the discussion, the students will be given a debriefing form about the whole purpose of the study.

6. Indicate the potential contributions of the study to your research area and/or the society.

It is thought that the results of this study will contribute to understanding the nature of the discussions taking place among students in EAP classrooms, thus informing teachers about the kind of scaffolding they should provide their students. It might also give some insight into the design of activities/materials to promote more critical thinking in class work discussions.

7. Write down the titles, dates of previous research projects you have conducted or that you have taken part in and the names of funding institution(s) if any.

Researcher's : Name-Surname: **Buket Esra Tarakçıoğlu** Signature _____

Supervisor's/Advisor's: Name-Surname: **JoDee Walters** Signature _____

*Some details may be different on the application forms as they were submitted before the study was conducted.

APPENDIX D: INFORMED CONSENT FORM

This study is conducted by Buket E. Tarakçıoğlu. The aim of the study is to collect data about the class discussions of participants. Participation in the study must be on a voluntary basis. No personal identification information is required. Your answers will be kept strictly confidential and evaluated only by the researcher; the obtained data will be used for scientific purposes.

The study does not contain anything that may cause discomfort in the participants. However, during participation, for any reason, if you feel uncomfortable, you are free to quit at any time.

At the end of the discussion, your questions related to the study will be answered. We would like to thank you in advance for your participation in this study. For further information about the study, you can contact Buket E. Tarakçıoğlu from the Department of Modern Languages (Room: S103 Tel: 210 3925; E-mail: tbuket@metu.edu.tr)

I am participating in this study totally on my own will and am aware that I can quit participating 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).

Name Surname Date Signature

----/----/-----

Gönülü Katılım Formu

Bu çalışma, Okt. Buket E. Tarakçıoğlu tarafından yürütülen bir çalışmadır. Çalışmanın amacı, öğrencilerin grup tartışmalarıyla ilgili bilgi toplamaktır. Çalışmaya katılım tamamiyle gönüllülük temelinde olmalıdır. Çalışmada sizden kimlik belirleyici hiçbir bilgi istenmemektedir. Cevaplarınız tamamiyle gizli tutulacak ve sadece araştırmacı tarafından değerlendirilecektir; elde edilecek bilgiler bilimsel yayımlarda kullanılacaktır.

Çalışma genel olarak kişisel rahatsızlık verecek herhangi bir durum içermemektedir. Ancak, katılım sırasında herhangi bir nedenden ötürü kendinizi rahatsız hissederseniz cevaplama işini yarıda bırakıp çıkmakta serbestsiniz. Çalışma sonunda sorularınız cevaplanacaktır. Bu çalışmaya katıldığınız için şimdiden teşekkür ederiz. Çalışma hakkında daha fazla bilgi almak için Modern Diller Bölümü öğretim elemanlarından Buket E. Tarakçıoğlu (tbuket@metu.edu.tr) ile iletişim kurabilirsiniz.

Bu çalışmaya tamamen gönüllü olarak katılıyorum. Verdiğim bilgilerin bilimsel amaçlı yayımlarda kullanılmasını kabul ediyorum. (Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).

İsim Soyad Tarih İmza

----/----/-----

APPENDIX E: DEBRIEFING FORM

This study, as stated before, is conducted by Buket E. Tarakçıoğlu, Department of Modern Languages, METU.

This study essentially investigates the discourse of critical thinking skills as reflected in the second language of EAP students. For this purpose, the amount of critical thinking and the quality of the language used will be examined.

It is aimed that the preliminary data from this study will be obtained at the end of April 2008. These data will be utilized only for research purposes. For further information, about the study and its results, you can refer to the following names. We would like to thank you for participating in this study.

Buket E. Tarakçıoğlu (Room: S103; Tel: 210 3925; E-mail: tbuket@metu.edu.tr)

KATILIM SONRASI BİLGİ FORMU

Bu çalışma daha önce de belirtildiği gibi ODTÜ Modern Diller Bölümü öğretim elemanlarından. Buket E. Tarakçıoğlu tarafından yürütülen bir çalışmadır. Bu çalışmada temel olarak akademik ingilizce öğrencilerinin eleştirel düşünce yeteneklerinin sınıfta ikinci dillerine nasıl yansıdığı (söylemi) araştırılacak ve incelenecektir. Bunun için ayrıntılı olarak eleştirel düşüncenin miktarına ve kullanılan dilin kalitesine bakılacaktır.

Bu çalışmadan alınacak ilk verilerin Nisan 2007 sonunda elde edilmesi amaçlanmaktadır. Elde edilen bilgiler sadece bilimsel araştırma ve yazılarda kullanılacaktır. Çalışmanın sonuçlarını öğrenmek ya da bu araştırma hakkında daha fazla bilgi almak için aşağıdaki isimlere başvurabilirsiniz. Bu araştırmaya katıldığınız için tekrar çok teşekkür ederiz.

Buket E. Tarakçıoğlu (Oda: S103; Tel: 210 3925; E-posta: tbuket@metu.edu.tr)

APPENDIX F: CRITICAL THINKING FRAMEWORK

Critical Thinking

1. Clarifying/defining (examples, defining terms, pointing at another aspect of the issue, metaphors)
2. Analysis/synthesis (similarities, differences)
3. Enhanced rephrasing (with additions)
4. Offering solutions/direction/ a course of action, (statements of should)
5. Inference/interpretation (reasoning, consequential)
6. Brief and triggering arguments or questions

Not Critical Thinking

1. Repetition without really adding anything new
2. Unclear/unfocused or irrelevant idea or examples
3. Accusational or defensive statements or questions without satisfactory explanation or evidence
4. Emotional statements that include feelings
5. Logical fallacies

APPENDIX G: EXAMPLE OF CODED TRANSCRIPTS

Public 4: /But I think you promised firstly that you will build up a bridge with public and celebrity And we give you stone and cement and then what I'm looking for, eee ... you built up only stone to each other. /unclear and you prevent us to go and see each other and share some of our feelings. / unclear because they don't they can't show their own feelings because of you. And because you ... eee ... change their feelings and you give us what you want to say and also ... eee / unclear you passed your lines and this means war and you declared the war (laughter). Innocent people will harm./ non-CT So we are defend ...we are trying to defend our children from your bad effects and also from celebrities bad effects, I have to say this, but you declared the war. / non-CT

Teacher: Yea

Paparazzi 2

46- /As I said before, we just chase the celebrities that wants us ... eee ... to chase. Eee there are celebrities that paparazzis ignore ... there are some celebrities ... that paparazzis just pass off and don't look at. And there are some celebrities paparazzis is..paparazzis ... are ... tail of them... (*Together with the one below*).

Paparazzi: focus

Paparazzi 2:

focus on them. Focus on them. Eee ... so ... eee ... celebrities want actually. / CT

Celebrity: cık

Paparazzi 2: / Some of them not all of them (laughter). We chase eee ... eee ... some of them. / non-CT

Teacher: Some of them.

Public 4:

47- /But you can't see ... you are ... eee ... harming our children. Maybe you are eee ... doing what they want but eee ... you affect our children's mind/ CT

and you cause a innoc ... you cause a unconscious...you cause creating unconscious people./ Undecided

Paparazzi 2: /eee I didn't say that what we are doing is ethical (laughter). But some public want us too. Most of public want use too. Most of the people want us. / Undecided

Paparazzi 5:

48- /We put smart icons on the programs. Children shouldn't watch them./ (laughter) CT

Public 4: /But you cannot … you should protect them. / nonCT

Paparazzi 5:

49- /YOU should protect your children. You should take them away from the screen and put them into bed./ (laughter) **CT**