

ASENA AKKAYA

EMOTIONAL PROSODY AND ACCENT PROCESSING:
A BILINGUAL PERSPECTIVE

Bilkent University 2023

EMOTIONAL PROSODY AND ACCENT PROCESSING: A BILINGUAL PERSPECTIVE

A Master's Thesis

by

ASENA AKKAYA

The Department of Psychology
İhsan Doğramacı Bilkent University
Ankara
May 2023

To my beloved family



EMOTIONAL PROSODY AND ACCENT PROCESSING:
A BILINGUAL PERSPECTIVE

The Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by

ASENA AKKAYA

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF ARTS IN PSYCHOLOGY

THE DEPARTMENT OF
PSYCHOLOGY
İHSAN DOĞRAMACI BİLKENT UNIVERSITY
ANKARA

May 2023

EMOTIONAL PROSODY AND ACCENT PROCESSING: A BILINGUAL PERSPECTIVE

By Asena Akkaya

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 Arts in Psychology.

Burcu Ayşen Ürgen
Advisor

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 Arts in Psychology.

Hande Ilgaz
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 Arts in Psychology.

Duygu Sarısoy
Examining Committee Member

Approval of the of Graduate School of Economics and Social Sciences

Refet Soykan Gürkaynak
Director

ABSTRACT

EMOTIONAL PROSODY AND ACCENT PROCESSING: A BILINGUAL PERSPECTIVE

Akkaya, Asena
M.A., in Department of Psychology
Advisor: Burcu Ayşen Ürgen
May 2023

People use different kinds of cues to process and understand spoken language in their daily lives. Understanding language is not solely based on the processing of grammar and words but also the processing of paralinguistic (non-verbal aspects of spoken language) cues, which have an important role in the ultimate comprehension of spoken language. Emotional prosody is one of the crucial paralinguistic cues and is the focus of the present thesis. It refers to verbal cues that convey different types of emotions based on changing pitch, timbre and tone of voice during speech. Different theories attempt to explain how emotional prosody and its related cues are cognitively processed during speech comprehension, such as the Universality Hypothesis, the In-group Advantage hypothesis and Dialect Theory. This thesis examined the processing of two different paralinguistic cues (emotional prosody and accent) and were discussed within the perspective of these theories. This study had 3 main research questions: (1) how do bilingual individuals differently process emotional prosody and different accents in their first (L1) and second (L2) languages? (2) How do different accents affect the emotional comprehension of L1 and L2 speakers? and (3) is the processing of vocal emotions universal or does it include culturally-specific aspects? Bilingual Turkish-English participants were tasked with listening to sentences in either their L1 (Turkish) or L2 (English) that varied in their emotional prosody (happy, sad, neutral tones of voice) and their accent (accented Turkish; accented English). Electroencephalography (EEG) recordings were taken to measure the P200 event-related potential response, which indexes automatic emotional processing and has been shown to vary between first and second languages. The results showed that Turkish-English bilingual individuals showed higher P200 amplitudes when they

heard emotionally vocalized sentences in their L1 compared to sentences vocalized with neutral tone of voice, regardless of the accent. However, in their L2, they showed higher P200 amplitudes for neutral sentences compared to emotionally vocalized sentences. Overall, these results propose that the nativelikeness of the stimuli is crucial to emotional prosody processing (in line with the In-Group Advantage hypothesis), such that we process emotional information more automatically in our L1, regardless of the speaker.

Keywords: accent processing, bilingualism, EEG, emotional prosody.



ÖZET

ÇİFT DİLLİ BAKIŞ AÇISIYLA DUYGUSAL PROZODİ VE AKSAN

Akkaya, Asena
Yüksek Lisans, Psikoloji Bölümü
Tez Danışmanı: Burcu Ayşen Ürgen
May 2023

İnsanlar günlük yaşamlarında konuşulan dili anlamak için farklı türden ipuçları kullanırlar. Dili anlamak, yalnızca gramer ve sözcüklerin işlenmesine değil, aynı zamanda konuşulan dilin ve iletilen mesajların anlaşılmasında önemli bir role sahip olan paralinguistik (konuşma dilinin sözel olmayan yönleri) ipuçlarının işlenmesine de dayanır. Duygusal prozodi, çok önemli paralinguistik ipuçlarından biridir ve bu tezin odak noktasıdır. Konuşma sırasında değişen perde, tını ve ses tonuna bağlı olarak farklı duygu türlerini ileten sözel ipuçlarını ifade eder. Evrensellik Hipotezi, Grup İçi Avantaj hipotezi ve Şive Teorisi gibi farklı teoriler, konuşmayı anlama sırasında duygusal prozodinin ve bununla ilgili ipuçlarının bilşsel olarak nasıl işlendiğini açıklamaya çalışır. Bu tez, iki farklı paralinguistik ipucunun (duygusal prozodi ve aksan) işlenmesini incelemiş ve konuyu bu teoriler perspektifinde tartışmıştır. Bu tezde sorulan 3 temel soru vardır. (1) Çift dilli bireylerin birinci (L1) ve ikinci (L2) dillerinde duygusal prozodiyi kullanmaları değişiklik gösterir mi? (2) Çift dilli bireyler farklı aksanları anadillerinde ve 2ikinci dillerinde farklı şekilde mi işlerler? (3) Duygusal prozodinin işlenmesi evrensel midir yoksa kültüre özgü belirli etkiler gözlemlenebilir mi? Çift dilli (Türkçe-İngilizce) katılımcılara, duygusal prozodileri (mutlu, üzgün, nötr ses tonları) ve aksanları (Türk konuşmacı; İngiliz konuşmacı) bakımından farklılık gösteren Türkçe veya İngilizce cümleler dinletildi. Otomatik duygusal işlemeyi indeksleyen ve, birinci ve ikinci dil arasında değişiklik gösterdiği bilinen P200 komponentini incelemek hedeflenerek elektroensefalografi (EEG) kayıtları alınmıştır. Sonuçlar, Türkçe-İngilizce çift dilli bireylerin, aksana bakılmaksızın, nötr ses tonuyla seslendirilen cümlelere kıyasla, anadillerinde duygusal olarak seslendirilen cümleler duyduklarında daha yüksek P200 genliği gösterdiğini buldu.

Bununla birlikte, katılımcılar ikinci dillerinde, duygusal olarak seslendirilen cümlelere kıyasla nötr cümleler için daha yüksek P200 genliği gösterdiler. Genel olarak, bu sonuçlar, uyaranların anadil benzerliğinin, duygusal prozodi işleme için önemli olduğunu (Grup İçi Avantaj hipoteziyle uyumlu olarak), konuşmacıdan bağımsız olarak, duygusal ses tonlarını anadilimizde otomatik olarak işlediğimizi gösterdi.

Anahtar sözcükler: aksan işleme, çift dillilik, duygusal prozodi, EEG.



ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my lovely advisor Dr. Alba Tuninetti for her immeasurable kindness, patience and guidance. Ever since the start of my Master's, her efforts and manner were always the primary guiding light for me to learn and discover. I will always feel lucky that I had the pleasure of being her student.

I would like to thank my co-advisor Dr. Burcu Ayşen Ürgen for her support and guidance.

I owe special thanks to Dr. Hande Ilgaz for her contributions to my academic journey. I learned a lot from her and she always drove me to look at things from new perspectives.

I am also grateful for Dr. Duygu Sarısoy for accepting to be my thesis committee member and their valuable feedback.

I cannot thank Dr. Tuğba Uzer Yıldız enough for always guiding me through the academic journey and always being there as someone I could look up to.

I'd like to acknowledge the support I have gotten from our research assistants Alara Akcan, Aybüke Karadağ, Berk Yüce, Canberk Tunakan. Special thanks to Nihal Gökçe for their contributions. I owe special thanks to Berrak Müftüoğlu, I learned a lot from her.

I'm extremely grateful to have Berfin Akça and Cansel İnan for their support and lovely friendship. I'd like thank Yiğit Sever for always being there with me and accompanying me on this journey.

Last but not least, I want to thank my mom and my dad for their unending support.

TABLE OF CONTENTS

1	Introduction	1
1.1	Prosody	1
1.2	Emotional Prosody	2
1.3	Theories of Emotional Prosody Perception	3
1.4	Recognition Rate and Saliency of Basic Emotions	8
1.5	Bilingualism and Emotional Prosody	10
1.6	Processing of Accented Speech	15
1.7	Behavioral Methods vs EEG	20
1.8	Emotional Prosody and EEG	21
1.9	Current Study: Aims and Hypothesis	23
2	Method	25
2.1	Materials	25
2.1.1	Stimuli	25

2.2	Norming Study I	26
2.2.1	Results	27
2.2.2	Auditory Stimuli	27
2.3	Norming Study II	28
2.3.1	Results	29
2.3.2	Acoustics	31
2.4	Main EEG Experiment	33
2.4.1	Participants	33
2.5	Design	34
2.5.1	Procedure	34
2.5.2	ERP Recording and Pre-processing	36
2.5.3	ERP Analyses	38
3	Results	41
3.1	Statistical Data Analysis	41
3.2	Results	42
3.2.1	P200 (170-230ms)	42
3.2.2	LPC (450-750ms)	55
4	Discussion	58

4.1	Limitations and Future Direction	65
4.2	Conclusion	67
A	Stimuli	87
B	Language History Questionnaire	97
C	EEG Pre-Screening Form	107



LIST OF TABLES

2.1	Statistical results of Norming Study I, for English sentences . . .	27
2.2	Statistical results of Norming Study I, for Turkish sentences . . .	27
2.3	Acoustic features of English stimuli	32
2.4	Acoustic features of Turkish stimuli	32
3.1	Repeated Measures ANOVA (P200).	42
3.2	Repeated Measures ANOVA (LPC)	56

LIST OF FIGURES

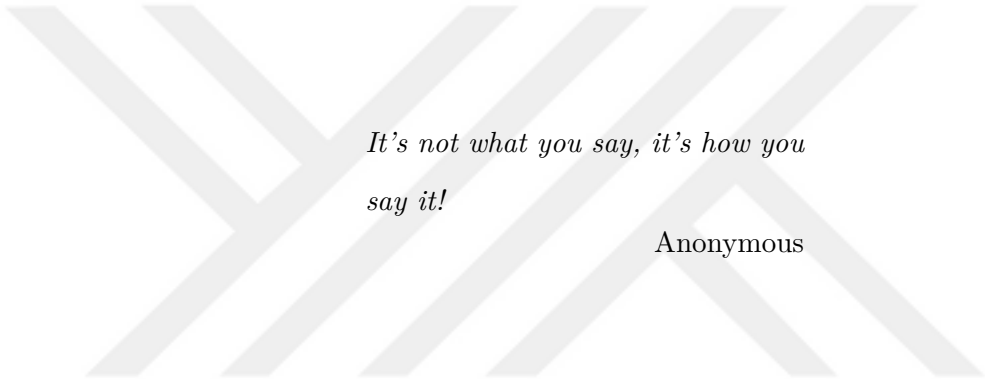
2.1	Variables for creating sentence categories.	30
2.2	Happily vocalized sentence: The top panel is an oscillogram and the bottom panel shows a spectrogram. The characteristic feature of a happy sentence is being louder, so it causes an envelope fluctuation and deviation from baseline in the oscillogram.	31
2.3	Vocalization of the sad sentence. The characteristic feature of a sad sentence is being quieter, so the waveform tapers out.	31
2.4	Neutral tone of voice. The waveform shows little to no fluctuation.	32
2.5	Participant with EEG cap placed prior to the experiment. The participant has given permission for the usage of their likeness in this Thesis.	36
2.6	Trial and timeline for the presentation of stimuli.	37
2.7	Regions of Interest (ROI)	39
3.1	Waveforms for Turkish sentences	50
3.2	Waveforms for English sentences	50
3.3	Waveforms for Turkish sentences of C3 and C4	52

3.4	Waveforms for English sentences of C3 and C4	52
3.5	P200 mean amplitudes for each emotion at frontal, central and posterior electrode sites. Error bars plot the standard error. . . .	53
3.6	LPC mean amplitudes for each emotion at frontal, central and posterior electrode sites. Error bars plot the standard error. . . .	54



Chapter 1

Introduction



*It's not what you say, it's how you
say it!*

Anonymous

1.1 Prosody

Language does not just consist of grammar and words. Instead, it consists of many components and each component contributes differently to the final overall meaning. According to Wilson and Wharton (2006), during communication, paralinguistic cues provide valuable insight about interlocutors' emotions and social intentions. Prosody is one of the important paralinguistic (non-verbal aspects of spoken language) components of this entity. It refers to the melodic and rhythmic features of spoken language. Speech prosody conveys two distinct sets of information: the speaker's emotional state (Fairbanks & Pronovost, 1938), called "emotional or affective" prosody (Monrad-Krohn, 1947), and cues related to syntax and pragmatics (Beach, 1991), called linguistic prosody (Monrad-Krohn, 1947). Even though these two types of prosody are distinct in their functions, they

share a set of acoustic cues such as changing tone of voice, pitch, timbre, and pauses during speech (Juslin & Laukka, 2003). For instance, in linguistic prosody, distinct intonation patterns are used to distinguish between questions and statements, such as a rising pitch at the end of a question versus a lowering pitch at the end of a statement. On the other hand, in emotional prosody, hearing loud and high-pitched speech might be the cue for the emotional state of the speaker (they are probably angry). According to the literature, linguistic and emotional prosody have distinct neurocognitive systems for processing (Witteman et al., 2011; Grandjean, 2021). In addition, Schirmer and Kotz (2006) show that emotional prosody is mainly processed in the right hemisphere while still being bilateral (Berckmoes & Vingerhoets, 2004) and linguistic prosody is mainly processed in the left hemisphere (Heyrani et al., 2021). Throughout this thesis, we will only concern ourselves with emotional prosody.

1.2 Emotional Prosody

Human social communication relies heavily on the interpretation and response to another person’s prosody and their ability to produce different prosody during speech as well. The ability to recognize emotions through voice is a crucial aspect of social cognition and human communication, it allows us to decode social interactions cues and respond appropriately while considering the context. According to Paulmann and Uskul (2014), listeners are able to recognize emotions thanks to the tone of voice of the speakers without paying attention to the semantics of sentences.

The “three-stage working model for the processing of emotional prosody” was proposed by Schirmer and Kotz (2006) to explain how humans perceive emotional prosody. The first stage involves identifying the sensory and acoustic features of the speech. In the second stage the emotional meanings are extracted from these

identified vocal cues. In the final stage called “conceptual processing” the emotional information is extracted and compared to other types of paralinguistic cues that also convey emotional information. This comparison is important for higher-level cognitive processes. In this stage, the acoustic features that are related to the knowledge of emotions are compared to those retained in long-term memory. The basic emotions and emotion-specific knowledge are stored separately in an associative memory network (Bower, 1981; Bower, 1987; Niedenthal et al., 1994). Pell (2005) asserts that when accessed in the auditory or visual modalities, the prototypes of acoustic or sensory elements linked to emotional expression could trigger the long-term memory representations stored and therefore ease vocal emotion recognition. Indeed, Paulmann et al. (2013) showed that emotional meanings in speech are embedded in vocal cues. They found that when participants heard the sentences vocalized with different emotional tones, they could correctly identify target emotions even if the sentences were made up from pseudowords. This is likely the effects of associations in memory over time and thanks to this, after initial processing, people easily retrieve information about intended emotions from previous representations in memory therefore they are able to access and process semantic information with this retrieved *information*.

1.3 Theories of Emotional Prosody Perception

Recognition of emotions is evolutionarily, socially, and biologically crucial for human communication because receiving and decoding important cues and reacting to them appropriately is vital for survival or social interactions in our daily lives. People continuously collect and process various cues from the environment. Bruce and Young (1986) claim that recognizing faces, identifying them, and being aware of the facial expressions to understand the emotional states of interlocutors’ are provides a lot of important information and increased the chance of successfully surviving. Bruce and Young proposed a functional model for face perception.

According to this model, facial expressions are categorically recognized during speech perception within 3 steps. First, analysis of facial information then analysis of affective facial expression information finally, analysis of face identity. Cowie and Cornelius (2003) show that the same appraisal process is valid for vocal emotional perception. In addition, Belin et al. (2004) consider speech as an *auditory face* that helps us to identify speakers' identity and emotional state. It creates an image that includes various information. Furthermore, they show that different brain areas are responsible for processing these various vocal information such as the sex of the speaker, the owner of the voice, or the emotional state of the speaker (e.g. anterior temporal lobes in both hemispheres for identification of speaker, right temporal lobe and right inferior prefrontal cortex for emotional prosody detection). Different theories of emotional perception and recognition have arisen based on various sets of data and how they have been interpreted.

According to Ekman (1994) and Izard (1994), basic emotions (happiness, sadness, fear, disgust, anger, surprise) should have distinct types of expressions in the face and the voice to ease coordination and the appraisal of emotions and communication. Research shows that vocalization of basic emotions are correctly identified when the sentences were spoken in different languages and listened to by different language speakers even if listeners have not exposed spoken languages before (Pell et al., 2009b; Scherer et al., 2001; Thompson & Balkwill, 2006).

Due to the biological significance of nonverbal expressions in social interactions and their role in adaptive behavior, these emotional expressions are thought to have certain unchanging characteristics that enable them to be identified without reliance on cultural or learned factors when presented through facial expressions (Ekman & Friesen, 1971) or vocalization (Scherer et al., 2001). Furthermore, according to Pell et al. (2009a), basic emotions are vocalized in similar manners across cultures in terms of acoustic properties. Acoustic parameters (mean, range and speech rate of emotional vocalizations) are responsible for similar variances across four different languages in given emotional expressions. This suggests

that speakers of English, German, Hindi, and Arabic utilize these features to a comparable degree in distinguishing a set of *basic* emotions. As a specific example, Pell (2001) found similar acoustic parameters for vocal expressions of sadness in different cultures. Furthermore, the authors argued that the expressions in question are slow in terms of their speech rates, are produced with a low f_0 mean, and have a limited f_0 range (f_0 is the fundamental frequency of speech, it shows average of oscillations for each second). Thus, all of these findings set forth the universality of the emotional prosody processing hypothesis (Ekman, 1993; Russell, 1994).

According to this hypothesis, emotional prosody is detected universally and emotions are transmitted similarly in different cultures and different languages. Scherer et al. (2001) called this ability of recognition of emotions across cultures “universality principles or inference rules” of basic emotions’ processing in different languages. Research repeatedly shows that listeners correctly recognize vocal emotional expressions in different languages (Beier & Zautra, 1972; Albas et al., 1976; Van Bezooijen et al., 1983; Graham et al., 2001; Scherer et al., 2001; Thompson & Balkwill, 2006; Bryant & Barrett, 2008; Pell et al., 2009b). For example, Kramer (1964) showed that emotionally vocalized English and Japanese sentences were easily recognized by American monolingual participants even though English and Japanese are typologically distant languages. In addition, Bryant and Barrett (2008) found that American English speakers successfully identify target emotions while they hear sentences in the Amazonian language. Moreover, people from remote villages in Bhutan (in the Eastern Himalayas) easily identify the correct emotions when they are listening to English even though they have not been exposed to English before (Cordaro et al., 2016).

On the other hand, the acquisition of different languages requires integrating new prosodic systems which adds new cognitive load to our perception (Besson et al., 2002), especially if the new language is typologically distant from the native language in terms of the pronunciation and lexicon. Van Bezooijen et

al. (1983) found in-group advantage while processing emotional prosody during speech. Hence, same language users are better at correctly classifying and identifying target emotions than people from different cultures (out-group). According to their research, native speakers outperform non-native speakers when they detect emotional prosody vocalized by their in-group members. In particular, they found that Dutch participants show better performance on recognition of vocal emotion in Dutch compared to Japanese and Taiwanese.

In addition, people from different cultures might make different attributions for the prosody. For instance, if someone speaks English in a flat and monotone pitch, L1 English speakers may interpret them as bored or unfriendly. However, this is not the case for native speakers of Russian, as a monotone pitch pattern is common in their language. So emotional prosody can be affected by culture and language. The in-group advantage hypothesis might be an appropriate explanation for these differences. As evidenced by Scherer et al. (2001), native speakers are better at recognizing emotional prosody in their native language than non-native listeners. Participants who were recruited from three different continents (to ensure variability between languages and cultures) heard 30 pseudo-word sentences that were vocalized with 5 different emotions by four German actors. Even though all emotions were correctly recognized by all participants, German participants outperformed in the emotion recognition task than participants from other countries (Scherer et al., 2001). Earlier, Beier and Zautra (1972) had indicated that American listeners are more successful in recognizing emotional sentences vocalized by American actors compared to Japanese and Polish actors. Likewise, the recent findings show that Chinese-English bilinguals easily identified vocal emotional English sentences when they were vocalized by Chinese speakers compared to English speakers (Yan et al., 2022).

Recognition of emotional cues during speech might cross the cultural barriers and be identified universally; however, there are subtle cultural variations that might cause difficulty to communicate effectively. Even though the majority of

the emotional cues are preserved across these barriers, some of the cues are inadvertently lost along the way. Researchers recognize both universal and cultural differences for the display and interpretation of emotions so some of them have an intermediate position (Fiske et al., 1998; Fridlund, 1994; Mesquita et al., 1997; Rosenthal et al., 1979). Elfenbein and Ambady (2003) have proposed a “Dialect Theory” to explain this in-group advantage and universality together (what I will term the intermediate position). They proposed this theory inspired by Tomkins and McCarter (1964) metaphor for recognition of emotions across cultures: Cultural differences in emotional expression are like *dialects* of the more universal grammar of emotion (p. 127; see also Rosenthal et al., 1979) (as cited by Elfenbein et al. (2007)). The general idea is that the recognition of emotions is akin to that of dialectal linguistic differences. Like different dialects within a language, how humans portray and perceive the ways of conducting emotions varies across groups. As such, people from the same culture are more prone to correctly understanding each other when conveying emotions (Elfenbein & Ambady, 2002; Juslin & Laukka, 2003). According to this hypothesis, the advantage of the in-group is reduced when groups have closer geographic proximity, typological proximity of their languages or greater cross-cultural interaction with one another. For example, Scherer et al. (2001) found that German participants show better performance recognizing vocal emotions in Dutch (German and Dutch are from the same language family) compared to Malay (German and Malay are typologically distant languages). Moreover, given some time, individuals seem to acquire the ability to comprehend the emotions of individuals from foreign cultures, as noted by Elfenbein and Ambady (2002) and Elfenbein and Ambady (2003). For instance, participants from Quebec (Canada) and Gabon (Africa) show cultural differences with the expression and identification of emotions (they used different strategies and cues for identification), but all participants correctly labeled target emotions after a short exposure to the emotional expression styles of out group members (Elfenbein et al., 2007). Likewise, Scherer and Wallbott (1994) revealed

significant evidence supporting both universality and cultural variations for emotion recognition with the data collected from 37 different countries, they found that all participants correctly recognized target emotions. However, they rapidly processed the stimuli that were vocalized by the speaker who has the same native language. Moreover, Thompson and Balkwill (2006) showed that although English, Chinese, Japanese, German, and Tagalog individuals correctly recognized the vocal emotions across languages, they still showed better performance in their native language.

Taken together, these findings suggest that the capacity to identify emotions conveyed through vocal expression is a universal capability, but it is also impacted by cultural and linguistic factors. These results support the idea that both biological factors and environmental influences may play a role in the ability to recognize emotions conveyed through speech.

1.4 Recognition Rate and Saliency of Basic Emotions

According to Ekman et al. (1969), basic emotions are universally recognized. However, in the emotion literature, while the recognition rate of basic emotions is found to be high, it varies between emotions. The vocalization of some emotions are identified better than others (Pell et al., 2009a; Thompson & Balkwill, 2006; Johnson et al., 1986; Scherer et al., 1991).

For example, anger and sadness are identified more rapidly and accurately than other emotions (Banse & Scherer, 1996; Johnson et al., 1986; Kramer, 1964; Murray & Arnott, 1993; Thompson & Balkwill, 2006; Wallbott & Scherer, 1986) while other research showed that fear is the most salient emotion, and it is recognized more easily than vocalization of other emotions (Levitt, 1964; Zuckerman et al., 1975). In addition, Pell and Kotz (2011) indicated that the shortest reaction

time interval was observed for recognition of fear and sadness from voice.

Likewise, Paulmann and Uskul (2014) claimed that regardless of the language, sentences vocalized with an angry tone of voice were recognized better than happy sentences. Ekman (1992) claimed that negative emotions might be evolutionarily more important than positive emotions. In terms of evolutionary point of view, the neurocognitive system gives priority to negative emotions that indicate danger and aggression, enabling individuals to respond quickly and appropriately to unwanted (vocal) stimuli. Similarly, Öhman and Mineka (2001) claimed that it may be more advantageous to recognize potential danger (anger, fear) than a safe and pleasant situation (e.g., happy, surprise). Differences in how individuals recognize different emotions may be due to biological or evolutionary factors. Different studies show that participants have some difficulties while recognizing a happy tone of voice compared to other basic emotions (Scherer et al., 2001; Levitt, 1964; Zuckerman et al., 1975; Paulmann & Pell, 2011). However, the accuracy in recognizing happiness is highest for the same group or culture members (in-group advantage) (Elfenbein & Ambady, 2002).

Some researchers also claim that positive emotions might be a vital element within a culture instead of across cultures because they ease group cohesion and communication. In addition, Bhattacharya et al. (2016) suggested that positive emotions might be used in a linguistically specific way. They claim that speakers would be much better at recognizing prosodic cues of positive emotions in their native language with the in-group advantage. Indeed, Paulmann and Uskul (2014) show that Chinese and British participants easily recognized a happy tone of voice when they heard in their native language, but it took more time when they heard sentences vocalized with happy tone of voice in foreign languages.

On the other hand, according to Sauter et al. (2010a), the existing literature is mostly based on negative emotions for emotional prosody studies. Research with positive emotions is highly neglected. Thus, the importance of working with

a balanced number of emotions from each category (negative-positive and also neutral (as a control variable)) should be highlighted in the future studies.

1.5 Bilingualism and Emotional Prosody

Languages that we have shape our perception and our way of seeing the world. It affects our appraisal, thoughts, and cognition (Lindquist et al., 2015). Having one or more than one language triggers important questions like whether humans see the world from the same window in their first (L1) and second languages (L2) or whether people who are monolingual, bilingual or multilingual are prone to use similar cues to figure out spoken languages.

Research has shown that there are differences in the way native speakers and L2 learners process language. Some studies have focused on similarity between L1 and L2 speakers' for syntactic information usage (Deniz, 2022), while others suggest that L2 speakers are more prone to use semantic information than L1 speakers (Grüter et al., 2020). However, the explanation for this discrepancy is still debated. According to the Shallow Structure Hypothesis (SSH), L2 speakers are prone to rely on semantic and pragmatic cues; they are more sensitive to non-grammatical cues than L1 speakers (Clahsen & Felser, 2006). Similarly, Cunnings (2017) showed that native (L1) and non-native (L2) speakers used different cues during processing of complex written sentences.

The discrepancy in cue usage for semantic, syntactic and pragmatic information that is mentioned above triggers an intriguing question for this thesis: how are paralinguistic cues (especially vocal) differentially weighed in L1 and L2? For the recognition of vocal emotions, one of the most important points is whether the ability to recognize and process emotions through prosodic cues is the same in a speakers' first language (L1) and second language (L2). Previous research on L2 learners' emotion recognition ability in their L2 has produced mixed results.

These studies typically involved identifying different emotions produced by emotional prosody using forced-choice tasks. In studies that compared L2 learners with native listeners, Graham et al. (2001) and Altrov (2013) showed that L2 learners had more difficulty recognizing emotions in their L2 compared to native listeners. On the other hand, Dromey et al. (2005) showed that L2 learners performed similarly to monolingual native listeners in the emotion recognition task in their L2 but polyglot native listeners outperformed both groups for the same language condition. The authors claim that polyglots are prone to having additional decoding skills in their native language for emotion recognition.

Contrarily, Wei et al. (2022) found that L2 speakers are better than native speakers at the detection of acoustic differences between the stimuli vocalized with different emotional tones and Zhu (2013) indicated that L2 learners actually outperformed monolingual listeners for recognition of emotions. However, in Zhu (2013), they conducted the research with Mandarin-Dutch bilinguals and Mandarin monolinguals; Mandarin uses tonal information both grammatically and paralinguistically. Thus, they suggested that acquisition of non-tonal language (Dutch) may have an advantage in focusing on the prosodic aspects of tonal information for vocal emotion recognition.

Importantly, as mentioned above, all of these studies were conducted with forced-choice tasks. There may be some drawbacks to using this answer format. According to Russell (1993), it is unclear what conclusions can be obtained from using closed-answer tasks for emotion recognition. When there are many choices, the subject makes a relative judgment, picking the one that seems most plausible. He claimed that the most likely response option does not always represent the response that the subject truly meant. In addition, the proficiency levels of L2 speakers' might be varied between studies and this situation might be another cause of this dichotomy in the results.

According to Altmann et al. (2012), language processing is affected by the

emotional content of language down to a single word. This claim is backed up by further research on monolingual participants which show faster reaction time for the processing of emotionally valenced words compared to neutral words (Kensinger & Corkin, 2003). On the other hand, as seen in many linguistic studies, we also observe in the emotion processing literature that studies have been conducted with mostly monolingual English participants and the results have been attempted to be generalized. Pavlenko (2012) also highlighted this situation, as she claimed that much of the research on emotional language has been carried out with monolingual English speakers; as such, interest in the relationship between emotions and bilingualism should be amplified to discover general trends in emotion processing literature.

In the literature, as in other cognitive aspects, a bilingual advantage for emotion recognition might be the case as well (Luk et al., 2011; van den Noort et al., 2019; Paap et al., 2015). Being bilingual can offer advantages when it comes to expressing emotions. The ability to speak two languages can benefit individuals in expressing their emotions, as it provides them with a broader range of vocabulary and cultural contexts that are unique to each language (Freeman et al., 2016). Numerous studies have shown that individuals who grow up in a bilingual environment become more sensitive and capable of interpreting a speaker's communicative intent than their monolingual counterparts. Bilingual individuals frequently monitor the speech and nonverbal cues to understand spoken language and determine strategies to respond appropriately (Ben-Zeev, 1977; Comeau et al., 2003; Comeau et al., 2007; Cummins & Mulcahy, 1978; Genesee et al., 1975; Siegal et al., 2009; Yow & Markman, 2007; Yow & Markman, 2009) (as cited by Yow and Markman (2011)). Thus, bilingual individuals might be more prone to use paralinguistic cues to get and convey emotions. Specifically, Yow and Markman (2011) found that bilingual children are more adept at using nonverbal cues to figure out an interlocutor's emotions than monolingual children, particularly in cases where there is a conflict between prosody and the meaning of speech.

Moreover, they showed that bilingual children were better than monolingual children at utilizing tone of voice and recognition of emotions when the meaning and vocalization are congruent as well. In addition, Morton and Trehub (2001) indicated that when intonation and content congruent, even monolingual children use vocal cues to detect a speaker's emotions effectively instead of solely relying on semantic cues.

The studies done by Dewaele (2010) and Pavlenko (2005) show that bilingual individuals easily recognize and process different emotions in their first language while they have some difficulties expressing and interpreting emotions in their L2. Altarriba and Basnight-Brown (2011), Dewaele (2004), and Pavlenko (2008) have all shown that emotion plays an increased role in the mother tongue (L1) of an individual compared to the languages they learn later in life (Lx). For example, recent research shows that reading, speaking or listening in L2 decreased the emotional intensity of contents (Weimer et al., 2022). Similarly, Ferré et al. (2022), showed that bilingual individuals rated emotional stimuli as “less emotionally intense” in their L2 compared to their L1. However, these studies were mostly conducted with late bilinguals (bilinguals who have learned a second language later in life, usually in adulthood).

Questionnaire-based studies have indicated that bilingual individuals perceive their L1 as more emotional than their L2, regardless of their level of proficiency in each language (Dewaele, 2010; Pavlenko, 2004; Pavlenko, 2005). Specifically, Pavlenko (2005) shows bilingual participants reported that even if they know the exact meanings of emotional words that were presented in their L2, they could not experience the emotions as intense as in their L1. However, self-reports may be influenced by some confounding factors (Degner et al., 2012). On the other hand, using more objective measures studies, such as reaction times (RTs), indicated mixed results. Anooshian and Hertel (1994) showed an advantage during the processing of emotional words in participants' L1 while others showed processing advantage in L2 (Ayçiçeği & Harris, 2004), and still others demonstrating no

difference between languages (Ayçiçeği-Dinn & Caldwell-Harris, 2009; Eilola et al., 2007; Ferré et al., 2013; Sutton et al., 2007). Additionally, EEG research utilizing a lexical decision task has revealed lower or delayed automatic emotional processing in participants' L2 (Conrad et al., 2011; Opitz & Degner, 2012) (as cited by Ponari et al. (2015)).

Another important aspect for processing emotion is exposure rate and proficiency level. In the literature, the studies comparing groups with and without L2 experience have often yielded mixed results due to confounding factors. For example, Altrov (2013) showed that listeners with L2 experience performed better than participants who did not have L2 experience, but it was difficult to know whether this is the influence of cultural or linguistic experience since the groups lived in different countries. Similarly, Shochi et al. (2010) found an advantage for L2 learners, but these participants had received formal training on recognizing emotions in their L2. Hence, it is difficult to properly account for the influence of language and training separately. Nevertheless, Zhu (2013) found that language experience can improve emotion recognition ability from tone of voice, the result shows that L2 speakers who have knowledge of the spoken language outperformed non-native speakers. However, another study found no difference in emotion recognition abilities of high and low proficiency L2 speakers (Graham et al., 2001).

Conversely, in research conducted by Bhatara et al. (2016), the authors revealed that being proficient in L2 may hinder the ability to identify emotional intonation in that language. Specifically, the study found that individuals with higher proficiency in their second language had lower accuracy in recognizing emotional tones in their second language. On the other hand, it is important to note that different emotions may differ in their recognition rate and saliency as highlighted before. In the literature, the conflicting results were found in studies that focused on perception across multiple emotions. A meta-analysis showed that through studies examining the perception of emotion from vocalization, anger

and sadness were the most consistently recognized emotions across different cultures. Conversely, happiness had the lowest cross-cultural recognition rate, but it is the emotion that had the highest in-group advantage (other positive emotions were not studied enough to be included in the analysis) (Elfenbein & Ambady, 2002). Bhatara et al. (2016) claimed that positive emotions might be transmitted with a culturally and linguistically specific way. This means native speakers have an advantage for transmission and interpretation of positive emotions and usage of prosodic cues compared to L2 speakers. These findings consolidate the view of why using balanced emotion (using equal numbers of emotions from each category e.g. positive and negative) is important, and crucial to examine how the same individual perceives emotional prosody in both their native and second language, rather than comparing independent and different groups of native and non-native participants. Indeed, the literature lacks studies that examine the same participants' emotional prosody processing abilities in their L1 and L2 at the same time.

1.6 Processing of Accented Speech

Another crucial factor for the recognition of emotions and detection of paralinguistic cues of language is the accent. Just as the effective usage of emotional cues eases human communication, learning quickly how to calibrate and process accents is one of the vital requirements of effective communication (Cristia et al., 2012). In today's world with the effect of multiculturalism and globalization, people from many parts of the world are regularly exposed to various accents. A foreign accent is one of the noticeable outputs of non-native speakers (Flege, 1995; Gluszek & Dovidio, 2010), and it creates a challenge for listeners. According to Wells and Wells (1982), having an accent means that a speaker has phonological, lexical, and prosodic deviations from native speech. Similarly, Clarke and Garrett (2004) suggested that accents can be placed on a perceptual scale based

on how different they sound and how deviant they are from native speech, what they term the Perceptual Distance Hypothesis. This scale can be used to rank foreign accents at the far end of the spectrum and regional accents somewhere in between. Results from Munro and Derwing (1995) supported this hypothesis; native speakers find it difficult to understand foreign-accented speech, which can result in comprehension and processing difficulties. In addition, Floccia et al. (2006) found that words spoken in a regional accent were identified more slowly than in a home accent (a delay of 30 ms) and even more slowly in a foreign accent (a delay of 100 to 150 ms).

Hatzidaki et al. (2015) show that foreign-accented speech interferes with the processing of positive emotions while this effect was not observed for native accented speech. This might be related to the positive emotions that are conveyed to culture and linguistic specific ways (in-group advantage). In the literature, research shows that listeners prefer native-accented speech over foreign-accented speech for processing. Specifically, most research indicates that adults are prone to more positively categorize (in terms of professional success, social status, and level of education) interlocutors who have native accents compared to foreign accents (Ryan, 1983; Giles & Billings, 2004; Lev-Ari & Keysar, 2010; Fuertes et al., 2012; Pantos & Perkins, 2013). In addition, parallel results have been reported for children, they show a preference toward native-accented peers as playmates (Kinzler et al., 2009; Souza et al., 2013). Moreover, McDonald et al. (2018) examined the effects of different types of language experience on speech processing and comprehension. They conducted the research with bilingual (English-Spanish) and monolingual (English) children; children from both groups responded slower when they heard foreign accented speech (English sentences vocalized by Spanish native speakers).

On the other hand, regardless of being L1 or L2 speakers, people encounter accented speech in their daily lives. Contemori and Tortajada (2020) show that L1 speakers show lower attention to the intent and emotions of speakers who have a

foreign accent. It can be argued that processing of foreign-accented speech is even more challenging for second-language (L2) listeners than native (L1) listeners. This is not only because of the variety between canonical and non-canonical speech, but also because comprehension occurs in the listeners' L2.

Research indicates that comprehending native-accented speech in L2 is already a difficult task. For instance, the perception of L2 phonemes remains persistently inaccurate, and it is influenced by the phonemic inventory of the listener's L1 (Strange, 1995). However, the challenge of comprehending speech in L2 can be lessened when L2 listeners have heard foreign-accented speech which is congruent with their native languages. This phenomenon has been called the "interlanguage speech intelligibility benefit" (Bent & Bradlow, 2003; Hayes-Harb et al., 2008; Imai et al., 2005).

For example, Pinet et al. (2011) found that French-English bilinguals easily processed the English sentences when they were voiced by native French speakers (this condition is called L1 congruent foreign accented speech) compared to other foreign accented sentences. However, foreign accented speech (L1 congruent) in L2 is not always the sign for intelligibility between interlocutors (Stibbard & Lee, 2006). For instance, Major et al. (2002) conducted a study with Spanish, Chinese, Japanese, and English native listeners using English sentences vocalized by native speakers of these four languages. The results show that only native Spanish participants showed a processing advantage compared to other groups. Likewise, Munro et al. (2006) found that only the participants who are Japanese-English bilinguals show processing advantage for English sentences vocalized by an L1 speaker, compared to both Cantonese-English and Mandarin-English bilingual groups. However, the important point should be highlighted for this research. Only Japanese is a non-tonal language between these languages. Mandarin and Chinese are tonal languages which means the meaning of words might vary between tonations. Thus, the availability of accent and tonal information coming from tonal languages may interfere with each other. On the other hand, Major

et al. (2002) and Munro et al. (2006) found different results for native Japanese speakers as mentioned above. These mixed results might be obtained from recruiting participants in different proficiency or using different methods. Moreover, their claim (L1 congruent foreign accented speech in L2 is not always the sign for intelligibility) contradicts with the interlanguage speech intelligibility benefit hypothesis for different languages. This dichotomy in the literature should be explored more.

According to the Speech Learning Model (SLM) proposed by Flege (1995), the first language (L1) and second language (L2) of bilinguals share representational space in the brain and representations for L1 and L2 continuously interact with each other. Age of acquisition of L2 is an important factor that detects the interaction between L1 and L2 phonology. Most studies in the literature, have examined how L2 listeners process native-accented speech (i.e. English-Dutch bilinguals heard Dutch sentences vocalized by Dutch speaker) and L1-congruent foreign-accented conditions (i.e. Dutch sentences vocalized by English speakers) and have shown processing of accented. However, there is a need for more research on direct comparison between native accented speech and foreign accented (L1 congruent) speech for both languages of bilingual individuals to see processing differences for L1 and L2 and the effects of accent. Especially for late bilinguals who acquire L2 within a limited context (i.e. only for academic purposes), they might be biased to standard native accented speech (what they have been exposed to during initial acquisition).

Therefore, considering different types of bilingualism (e.g., simultaneous [learned at the same time] vs sequential [learned one language after the other] vs late [learned an L2 in adulthood]) is one of the important factors in understanding the effects of accent in speech perception. On the other hand, Grey et al. (2019) found that there is a difference in how native and foreign-accented speech is processed grammatically by L2 speakers while they found that L2 speakers show the same semantic processing (related to understanding the meaning of

speech) for native and foreign accented speech. However, the current literature lacks an answer for whether L1 and L2 speakers have different perceptual mechanisms between differently-accented speech for emotional language processing. In the literature, few studies have focused on different accents and processing of emotions together (Hatzidaki et al., 2015; Sun et al., 2022). Sun et al. (2022) examined how accentuation affects emotional word processing, and their task was based on understanding the semantics of emotional words. The results suggested that accented speech can have a significant impact on the emotional content of spoken language, both in implicit and explicit processing. Accented words ease access to the meanings of emotional words compared to neutral words.

Similarly, Hatzidaki et al. (2015) used emotionally-laden words such as “gift” (happy or positive connotation) and they examined whether foreign accented speech affects higher order (semantic) processing for recognition of emotion-laden words. Specifically, they investigated how native Spanish listeners identify affective words that are vocalized with different accents (four foreign accents and one native accent) in their native language. They claimed that processing of emotional words requires phonological awareness for familiarity of the words. So the accent might be crucial for processing emotions. They found that emotional words were easily processed compared to neutral words without significant differences between native and foreign accented conditions. However, foreign accented speech causes a negativity bias, such that participants respond more quickly to negative than positive words when they are vocalized with foreign accents. Thus, these two studies prove that accent has a crucial role on the processing of emotional contents, it affects how we perceive and react to emotions. On the other hand, these studies only focus on emotionally valenced words, not the effects of emotional prosody and accents together, and additionally, both of these studies were conducted with monolingual individuals. So, again we do not know how bilingual individuals differently process vocal emotions and accents in their L1 and L2. These findings trigger another important question: how do two different paralinguistic cues (emotional prosody and accent) affect each other?

1.7 Behavioral Methods vs EEG

To date, no study has looked at all three variables together: emotional prosody perception, accent, and L1 vs L2 status. To examine these variables, this study will use event-related potentials (ERPs) as measured by electroencephalography (EEG) (measuring the brain's electrical activity from the scalp). As mentioned earlier, most of the studies in the literature have been behavioral work such that it asks participants for their overt responses to emotional stimuli. However, overt responses can be influenced by factors that cannot be controlled by the experimenter (e.g. motivation, attention, environment). For instance, behavioral measures can cause observer bias; observers might be prone to evaluate the results as congruent with their hypothesis and expectations (Rosenthal, 1963). Moreover, the presence of the observer may cause different reactions from the participant knowing that they are being watched, or social desirability might affect their responses as well (Moskowitz, 1986; Kagan, 2007).

Another crucial drawback is the limited ability to assess nonverbal responses. For instance, behavioral measures are not enough to explain how fast our brain processes nonverbal cues as we only see the end result of processing, the overt response. These kinds of nonverbal cues can be important in conveying and processing emotional information. While EEG does not necessarily get rid of those factors, it does provide a more real-time fine-grained measure of processing and perception as it unfolds. It has several advantages over traditional behavioral methods in emotional prosody research. First of all, it provides a direct measurement of neural activity associated with emotional processing. Unlike behavioral methods which rely on self-report or indirect measures of emotional response, EEG can directly detect changes in brain activity associated with emotional processing, providing a more accurate and objective measure of emotional response. Furthermore, EEG has a high temporal resolution, which means it can detect changes in brain activity within the millisecond. This is important in research

that examines the processing of paralinguistic cues because their effects can occur rapidly and vanish rapidly again. Moreover, EEG can capture these rapid changes in emotional processing that might be missed with slower behavioral methods such as pressing the button for reaction time. Finally, there have been well-documented waveforms and electrical components that correspond to emotional linguistic processing in the brain - these are certain event-related potentials (ERP) components such as the P200, the N400 and the LPC.

1.8 Emotional Prosody and EEG

In the emotion processing literature, most research conducted with EEG has examined different facets of emotional processing. For example, Ashley et al. (2004) found that participants correctly identify the target emotions from facial expressions within the first 200 ms. Similarly, Schirmer and Kotz (2006) found that the first differentiation of vocal emotional sentences from neutral sentences starts within the first 200 ms (i.e. neurally, participants can differentiate between emotional vs neutral sentences within the first 200 ms, leading to later responses in identification or recognition). Thus, they identified the P200 component as an early sign of emotional prosody processing in speech. The P200 effect has been observed in later studies as well (Paulmann et al., 2008; Paulmann & Kotz, 2008; Sauter et al., 2010b; Liu et al., 2012) and show that the recognition of emotional voices is faster and more automatic compared to neutral voices. In addition to the P200 component, other ERP components have been used in emotional prosody studies. The early P200 component is believed to indicate increased attention to emotional stimuli to facilitate preferential processing, if necessary. Later ERP components are often associated with more profound processing processes (such as meaning assessment and access to schemas and representations in memory). In particular, late emotional prosody effects have been noted in a number of late ERP components, including the P300 (Wambacq &

Jerger, 2004), N300 (Bostanov & Kotchoubey, 2004), the N400 (Schirmer et al., 2002; Schirmer et al., 2005; Schirmer & Kotz, 2003; Paulmann & Pell, 2011), and a late positive complex (LPC) (Kanske & Kotz, 2007; Schirmer et al., 2013) (as cited by Paulmann et al. (2013)). These components were observed depending on stimuli, tasks, and experimental designs that were assigned. For instance, Paulmann et al. (2013) conducted a study that examined valence and arousal effects for emotional prosody. They found that the sentences which were vocalized with highly aroused tone of voice elicited a stronger P200 response compared to those that were elicited with a neutral tone of voice. Later, between 400 and 750 ms, a late positive complex (LPC) response was observed for emotional vocalizations compared to neutral. Again, their results supported the initial differentiation of vocal emotion is seen in the P200. Similarly, Chen et al. (2011) observed the P200 and the LPC during the processing of emotional prosody. The P200 was seen during decoding of acoustical cues between 150-250 ms after stimuli onset and the LPC was observed during the integration of vocal emotions and meaning of the sentences. Furthermore, recent work done by Weimer et al. (2022) showed that the LPC component was seen during vocal emotional processing for both native and L2 speakers (native German and Chinese-German bilinguals). The results show that native and L2 German speakers similarly process the emotional prosody and the meaning of German sentences.

Paulmann and Pell (2011) also showed that vocalization of basic emotions (anger, fear, disgust, sadness, surprise, happiness) were differentially processed from each other and neutral sentences within 200 ms. Moreover, Liu et al. (2012) conducted a study with non-verbal utterances vocalized with angry, happy, and neutral voices. They found that angry sounds were processed earlier (50 ms) than both happy and neutral, while happy and angry sounds showed increased P200 components. Similarly, Spreckelmeyer et al. (2006) show that the P200 component is more salient for happy voices than sad and neutral. As a result, the literature shows that the P200 component is an early sign for emotional prosody processing but the processing rate of basic emotions varies between them.

1.9 Current Study: Aims and Hypothesis

As highlighted before, in the current literature (to our knowledge) no prior study exists that investigates effects of L1 and L2 and accents together for emotional prosody processing. This thesis aimed to fill this gap by answering the following questions:

- How do bilingual individuals differentially weigh emotional prosody as a cue during speech comprehension in their L1 and L2?
- How do different accents affect the emotion comprehension of L1 and L2 speakers?
- Is processing of vocal emotions universal or does it include culturally-specific aspects?

We hypothesized that L2 speakers are more prone to use emotional prosody to attend to speech (based on Paulmann and Kotz (2008)) because L2 speakers need to pay more attention to other cues in order to arrive at the correct semantic interpretation while native speakers easily understand the speech without the need to process extra cues. We expect to see a larger amplitude in the P200 component of L2 speakers than L1 speakers when they hear emotional stimuli. In addition, when it comes to the LPC component, we expect to observe LPC effect for both groups as a result of integration of meaning and emotional tones in later ms.

On the other hand, accents would change the detection of emotional prosody, and as such, we predicted an interaction with accent. During the experiment, listeners heard English or Turkish sentences vocalized by both native Turkish and American English speakers. Therefore, there were Turkish statements vocalized in a native Turkish accent and in an accented American-English accent

and respectively similar for English sentences. Thus, we expect that for Turkish stimuli, L1 Turkish speakers more easily process native accents than foreign accents in Turkish. However, this situation would be changed for foreign accented speech with the effects of context of language that they acquired. All of the participants are late bilinguals, and they acquired their second language (L2) within school settings. We expected that when the participants heard the foreign accented speech (English sentences vocalized by an American English speaker), they would process those more easily than in their native accents.



Chapter 2

Method

In this section, I will cover the two norming studies done to design and finalize the stimuli used in the final EEG experiment, as well as the methods for the main EEG experiment. As such, I will focus on the stimuli first and detail the norming studies, then detail the participants, design, and methods of the main EEG experiment.

2.1 Materials

2.1.1 Stimuli

During this experiment, we aimed to transmit emotions only through tone of voice without any emotional meaning attributions that come from semantics. To achieve this, we created semantically neutral sentences in both Turkish and English. Our sentences were created with the words selected from OpusCorpus (<https://opus.nlpl.eu/>) such that sentences in both Turkish and English were semantically the same. OpusCorpus is a platform that includes various open source corpora in different languages. We used a parallel corpus that showed the frequency rates of words for each language (i.e. how frequent each word is in each

language separately) and were able to select words used at similar frequencies in both Turkish and English. 300 sentences were created by the author of this thesis in each language for a total of 600 sentences using the following criteria: the sentences should include at least 5 words and no more than 12 words. In addition, the number of verbs and nouns should be as close as possible for a sentence. After sentence creation, we conducted our first norming study to ensure that these sentences were indeed semantically neutral (that is, they did not elicit any emotionally valent response).

2.2 Norming Study I

For this norming study, 105 (42 Native English, 63 Native Turkish) participants were recruited. Participants who joined the norming studies were asked not to attend the main experiment. Participants rated only the stimuli in their native language. The norming study was released via a Qualtrics survey. On the start page of the experiment, there was information about the experiment and a consent form. Participants could terminate the survey at any time without any restrictions. The survey included three questions to judge each sentence based on Göz et al. (2017). As such, participants were asked to judge the created sentences in terms of neutrality, concreteness, and imageability on a 7-point Likert scale.

- How “neutral” is the sentence above? (“1” corresponds to no emotional response, “7” corresponds to a strong emotional response)
- How “concrete” is the sentence above? (“1” corresponds to very abstract concepts, “7” corresponds to very real and tangible concepts).
- How “imageable” is the sentence above? (“1” corresponds to not being able to picture this statement in your mind; “7” corresponds to being easily able to picture this statement in your mind).

2.2.1 Results

In general, participants understood the questions and rated the sentences as expected. For example, “This bottle is made of glass.” was rated as very imageable, whereas “This is gluten free apricot jam.” was rated as not very imageable. For the questions related to concreteness and imageability, we excluded sentences that received lower than a 5 out of 7 because we wanted sentences that were very concrete and very imageable. For neutrality, we excluded sentences that received higher than 3 out of 7 because we wanted sentences that were very neutral.

Table 2.1: Statistical results of Norming Study I, for English sentences

Descriptive	Neutrality	Concreteness	Imageability
Average	1.43	6.04	5.96
Standard Deviation	0.42	0.52	0.56

Table 2.2: Statistical results of Norming Study I, for Turkish sentences

Descriptive	Neutrality	Concreteness	Imageability
Average	2.16	6.31	6.16
Standard Deviation	0.47	0.39	0.44

Ultimately, 100 sentences were removed from the stimuli list for each language based on the above criteria, leaving 200 sentences for our auditory stimuli for the EEG experiment.

2.2.2 Auditory Stimuli

The remaining 200 English and 200 Turkish sentences were vocalized by two female speakers. Speakers were intentionally selected as female because Schirmer et al. (2002) indicated that listeners easily detect emotional content when speech is vocalized by female speakers. For our study, we recruited 1 native Turkish

and 1 native American English speaker to vocalize both Turkish and English sentences. This way, we would have native speakers for each set of sentences as well as non-native speakers for the other set of sentences, eliciting accented language.

The Turkish sentences were vocalized by a native accent (native Turkish speaker) and a foreign accent (native American English speaker), and the English sentences were vocalized by a native accent (native American English speaker) and a foreign accent (native Turkish speaker). These two speakers vocalized these sentences with three different emotional tones: happy, sad, and neutral.

The stimuli were recorded at the Bilkent University Department of Communication and Design (COMD) studio. This studio is sound-attenuated and is appropriate for recording sensitive auditory stimuli (it is also used for the recordings of the Bilkent Symphony Orchestra and Bilkent Radio). The stimuli were recorded with a handheld voice recorder (ZOOM H4n Pro) that was connected to the microphone (Sennheiser ME66 SHOTGUN). After the recording sessions, to get the stimuli ready, Audacity and Praat programs were used to edit sentences. Recordings for each language were divided into sentences and durations were noted down for each of them. Also, silence was added at the beginning and at the end of each sentence so there would be similar silence before and after each sentence was uttered. Then, Norming Study II was conducted to select the final stimuli.

2.3 Norming Study II

The second norming study was conducted to measure whether the sentences vocalized by the speakers in three different emotion tones reflected the emotions we intended to elicit.

Similarly to Norming Study I, the survey was released via Qualtrics. As in Norming Study I, guidelines about the experiment and written consent were presented before participants started the survey. 102 different participants (54 native Turkish, 48 native English) completed this norming study. Each stimulus was judged only by native speakers of the stimuli language. Participants who participated in the norming study were instructed not to participate in the main EEG experiment.

The survey consisted of 10 blocks: the participants judged 20 recorded sentences in each block. First, there was a clip where they could play the audio recording, then they were presented with three questions attached to each recording.

Participants were asked to judge the recordings via the following questions:

1. What kind of emotion is portrayed in this sentence? (Participants were given three options: happy, sad, and neutral.)
2. How well does the sentence convey that emotion? (Participants were expected to judge the intensity of the emotion on a 5-point Likert-type scale: 1 = not at all, 2 = slightly, 3 = moderately, 4 = very well, 5 = completely).
3. Did you understand the sentence? This question was to ensure whether the recording itself (accented or not) was clear. Participants replied to this question with a simple yes or no response.

2.3.1 Results

In general, participants rated the auditory sentences as expected: sentences that were vocalized as happy were rated as happy 53% of the time; sentences that were vocalized as sad were rated as sad 82% of the time; and sentences that were

vocalized as neutral were rated as neutral 90% of the time. Based on participant responses, we selected sentences that were (a) congruent with the expected emotion (i.e. sentences that were rated as the expected emotion); (b) were rated between a 2 and 4 on the intensity scale; and (c) were rated as clear in the final question. 30 sentences from each category were selected as our final stimuli for the EEG experiment.

Language	Emotion	Accent
English	Neutral	Native
	Happy	
Turkish	Sad	Foreign

Figure 2.1: Variables for creating sentence categories.

To summarize, this led to the following number of sentences being chosen:

- 30 TUR-Happy-native accented
- 30 TUR-Sad-native accented
- 30 TUR-Neutral-native accented
- 30 TUR-Happy-Foreign accented
- 30 TUR-Sad-Foreign accented
- 30 TUR-Neutral-Foreign accented
- 30 ENG-Happy-native accented
- 30 ENG-Sad-native accented
- 30 ENG-Neutral-native accented
- 30 ENG-Happy-Foreign accented
- 30 ENG-Sad-Foreign accented and,

- 30 ENG-Neutral-Foreign accented.

For the final list of stimuli, see Appendix 1.

2.3.2 Acoustics

Acoustic analyses were conducted through Praat (Boersma & van Heuven, 2001) to see whether the vocal pattern (acoustic features) of sentences are congruent with the target emotions' features. You can see examples (in Figures 2.2; 2.3; 2.4) for happy, sad, and neutral vocalization respectively. Every figure depicts the “The dress is hung on a hanger.” sentence, vocalized with different emotions.

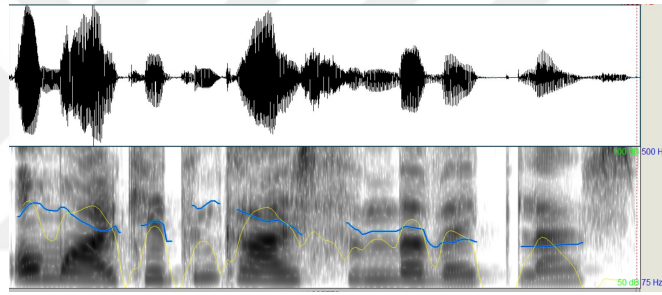


Figure 2.2: Happily vocalized sentence: The top panel is an oscillogram and the bottom panel shows a spectrogram. The characteristic feature of a happy sentence is being louder, so it causes an envelope fluctuation and deviation from baseline in the oscillogram.

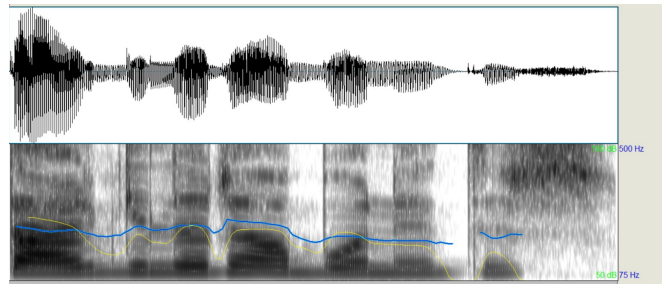


Figure 2.3: Vocalization of the sad sentence. The characteristic feature of a sad sentence is being quieter, so the waveform tapers out.

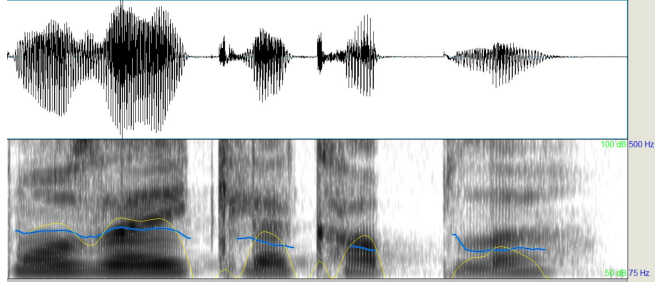


Figure 2.4: Neutral tone of voice. The waveform shows little to no fluctuation.

In addition, the acoustic features of all stimuli are presented in Table 2.3 for English stimuli and in Table 2.4 for Turkish stimuli.

Table 2.3: Acoustic features of English stimuli

Speaker	Parameter	Sad	Neutral	Happy
English	Average Duration	2.99 (0.60)	2.02 (0.37)	1.96 (0.43)
	Mean f0	322.28 (27.09)	267.18 (23.12)	320.32 (32.75)
Turkish	Average Duration	2.57 (0.45)	2.15 (0.61)	2.21 (0.46)
	Mean f0	234.97 (11.58)	235.86 (7.83)	282.45 (23.33)
Both	Average Duration	2.78 (0.56)	2.08 (0.51)	2.08 (0.46)
	Mean f0	278.63 (48.63)	251.52 (23.29)	301.38 (34.05)

Table 2.4: Acoustic features of Turkish stimuli

Speaker	Parameter	Sad	Neutral	Happy
English	Average Duration	3.01 (0.62)	2.46 (0.61)	2.54 (0.64)
	Mean f0	321.61 (26.16)	260.27 (14.53)	332.96 (40.82)
Turkish	Average Duration	2.16 (0.44)	2.07 (0.50)	1.91 (0.49)
	Mean f0	248.46 (16.09)	248.56 (12.18)	296.79 (19.61)
Both	Average Duration	2.59 (0.68)	2.27 (0.59)	2.23 (0.65)
	Mean f0	285.04 (42.71)	254.41 (14.54)	314.87 (36.61)

2.4 Main EEG Experiment

2.4.1 Participants

Sixty-one Turkish-English bilingual participants were recruited for the experiment from the surrounding community and universities. Data collected from 7 participants had to be excluded: two of them reported that they are left-handed and data from another 5 participants was also excluded due to noisy signals during recording. Thus, the data from the remaining fifty-four (32 female; 20 male; 2 decline to disclose) Turkish-English participants were used. The number of participants was determined in the light of earlier EEG studies in emotional prosody literature (i.e. Bhatara et al., 2016; Paulmann and Kotz, 2008; Pell and Kotz, 2011) as well as a power calculation. In addition, an equal number of participants were recruited to language conditions (27 English, 27, Turkish). Participants' age ranged from 19 to 33 years old ($M = 23.19$ years, $SD = 3.35$ years). All the participants were undergraduate or graduate students from Bilkent University and Middle East Technical University (METU). They had normal or corrected to normal vision and none of them reported hearing problems. Their written informed consent was collected and signed by each participant before the experiment.

The participants reported that they were late bilinguals: they started learning English in adolescence (average age: 13). Furthermore, participants were regularly exposed to English within their school settings from primary school onwards, including their undergraduate education.

The Bilkent University Ethics Committee fully approved the experiment including every material used and all experimental procedures. Participation of all individuals was paid financially (75 Turkish lira) or as 20 GE (general education courses) points. They were recruited through BAIS (Bilkent University Administrative Information System) mail, social media platforms (i.e. Twitter), and

snowball sampling technique which refers to people who had already taken part in the experiment suggesting it to their acquaintances who met the participation criteria. An EEG pre-screening form, demographic information form, and Language background questionnaire were completed online through Google Forms. Access to the files is restricted only to the PI, two research assistants from our lab, and the author of this thesis to protect participants' privacy.

2.5 Design

For the EEG study, a mixed design was used. We used between-participant design for languages (Turkish-English bilinguals heard only Turkish or English sentences during the experiment) while both emotion and accent conditions were within-participant conditions. This means each participant heard sentences vocalized with three emotions (happy, sad, neutral) and two accents (native-accented and foreign-accented). Every participant was randomly assigned to listen to only Turkish or only English sentences throughout the entire experiment. During the experiment, every voice recording was shuffled, and the participant heard the recordings pseudorandomly, such that each participant heard all six possible combinations of speaker and emotion. For instance, a participant hearing only English sentences for the experiment might hear Native English-Sad, Native English-Happy, Accented English-Sad, Native English-Neutral, Accented English-Neutral, Accented English-Happy, completing a set. Each participant heard a total of 180 sentences, which lasted approximately 25 minutes total.

2.5.1 Procedure

Prior to the experiment, participants were asked to complete a Language History Questionnaire (LHQ; Li et al., 2020; see Appendix B) and an EEG pre-screening form (Appendix C). The LHQ includes questions that measure participants'

proficiency level in the languages that they have. The EEG pre-screening form aims to detect participants' appropriateness for being participants in an EEG experiment. Once they arrived to the lab, they were welcomed in, given the consent form and explained the basics of the study. After answering any questions participants may have had, participants signed the consent form. After filling out the consent form, participants were invited to sit in a chair to have the EEG cap fitted for the recording (see Section 2.5.2: ERP Recording below for more detail). They were told that they would hear sentences and be asked simple identification questions about the sentences every once in a while. These questions were inserted to ensure that participants paid attention to the stimuli and were not distracted.

After cap fitting, participants were moved to the main stimulus presentation room, where they were seated in an immobilized chair to minimize muscle movement and noise that can impact EEG recordings. The distance was approximately 60 cm between the participants and the computer monitor and response keyboard. The experiment was designed via Python with the Psychtoolbox-3.0.17 version. Stimuli were presented via Etymotic (ER3) insert earphones (Lucid Hearing Holding Company, 2023) with 3M earbuds on a 20-inch monitor. A fixation cross was presented in the middle of the computer screen while the auditory stimuli played. To ensure that participants paid attention but did not overtly concern themselves with the emotional valence of the sentences, after the auditory stimulus was complete, we asked them to identify if a word had been present in the previous sentence. This probe word was shown on the screen and participants were asked to respond "yes" or "no" via mouse.

As such, throughout the experiment, the stimuli were presented in the following order:

- First, the fixation cross was presented within the first 500 ms.
- Then, participants heard auditory stimuli while continuing to see the fixation cross on the screen.



Figure 2.5: Participant with EEG cap placed prior to the experiment. The participant has given permission for the usage of their likeness in this Thesis.

- After the stimulus, a probe word popped up on the screen during the next 300 ms and vanished.
- Participants saw the “YES” and “NO” options to choose whether the probe word was in the stimulus or not. They were given 800 ms to respond.
- After their response, a blank screen was presented for 1500 ms before commencing the next trial.

2.5.2 ERP Recording and Pre-processing

During the electroencephalogram (EEG) session, 64 sintered (Ag/AgCl) MR-compatible elastic caps were used from the BrainCap MR system (Brain Products, GmbH). The electrodes are mounted on the caps according to the 10/20 system, which is a classification system developed by the American Electroencephalographic Society in 1991. The EEG activities were recorded via BrainAmp amplifiers and the Vision Recorder software (Brain Products, GmbH, Gilching, Germany). The elastic caps were carefully placed on the participants’ heads after

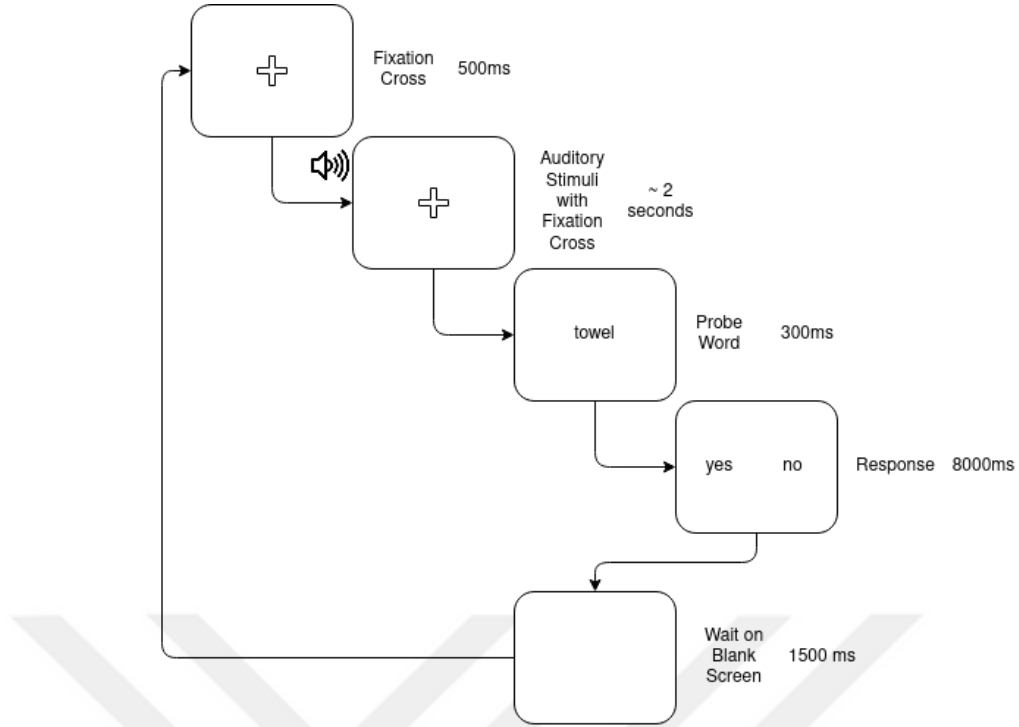


Figure 2.6: Trial and timeline for the presentation of stimuli.

receiving approval from the participants. The reference electrodes were selected as the FCz and AFz.

To increase conductivity, a special gel which is called ABRALYT 2000 FMS (Herrsching Breitbrunn, Germany) was applied with Q-tips and syringes. We aimed to decrease impedance and kept it under 10 k Ω . Electrocardiogram (ECG) was recorded for each participant as well: an ECG electrode was put on the back of the participant.

The Rigol DS1204B digital oscilloscope was used to adjust the delay between the auditory stimuli and EEG triggers to synchronize them. Due to hardware constraints, the EEG trigger is always sent earlier than the auditory stimuli without an added manual delay. For subsequent analysis, EEG data and triggers that we created to see the effects of stimuli were saved via BrainVision Recorder Software (Brain Products, GmbH).

To preprocess the EEG data Brain Vision Analyzer, version 2.1, and MATLAB were used. During the experiment, we could not get the signal from the F1

channel in both caps. Thus, F1 is eliminated in addition to the ECG during the preprocessing pipeline. Although there was no problem in removing it as the signals from the ECG would not be included in the analysis, F1 was interpolated after it was removed. Then rereferencing and resampling were applied. The sampling rate was decreased to 512 Hz. Then, an event list was created via ERPlab labeling our EEG triggers corresponding to each condition (language, accent, emotion).

Epochs were extracted from -200 ms pre-stimulus to 1000 ms post-stimulus. A high-pass filter of 0.1 Hz and a low-pass filter of 30 Hz was applied to exclude low and high wavelengths. Independent component analysis (ICA) was applied to detect eye blinks and other noisy activities. Then, unwanted eye movements were rejected, and fully automated artifact rejection was done. We used a threshold of 100 mV, and moving window 200 mV as the criterion for rejecting artifacts (above 100 mV was rejected). Data files with above 40% rejected trials were reviewed again to understand whether the problem was caused by a specific channel. If a specific channel was the cause (such as F1), they were interpolated for final analyses.

2.5.3 ERP Analyses

For ERP (P200 and LPC) analyses, 6 new channels were created by taking the average of 4 electrodes to form regions of interest (ROIs) (see Paulmann et al., 2013 for further information). We grouped electrodes according to regions of interest (ROI) for the ERP analysis. There were a total of six ROIs: Left frontal (LF): F5, F3, FC5, FC3; left central (LC): C5, C3, CP5, CP3; left posterior (LP): P5, P3, PO7, PO3; right frontal (RF): F6, F4, FC6, FC4; right central (RC): C6, C4, CP6, CP4; and right posterior (RP): P6, P4, PO8, PO4. ERPs were then plotted for these new channels. Based on previous studies, an early time window from 170 to 230ms (Paulmann et al., 2011) was considered for the

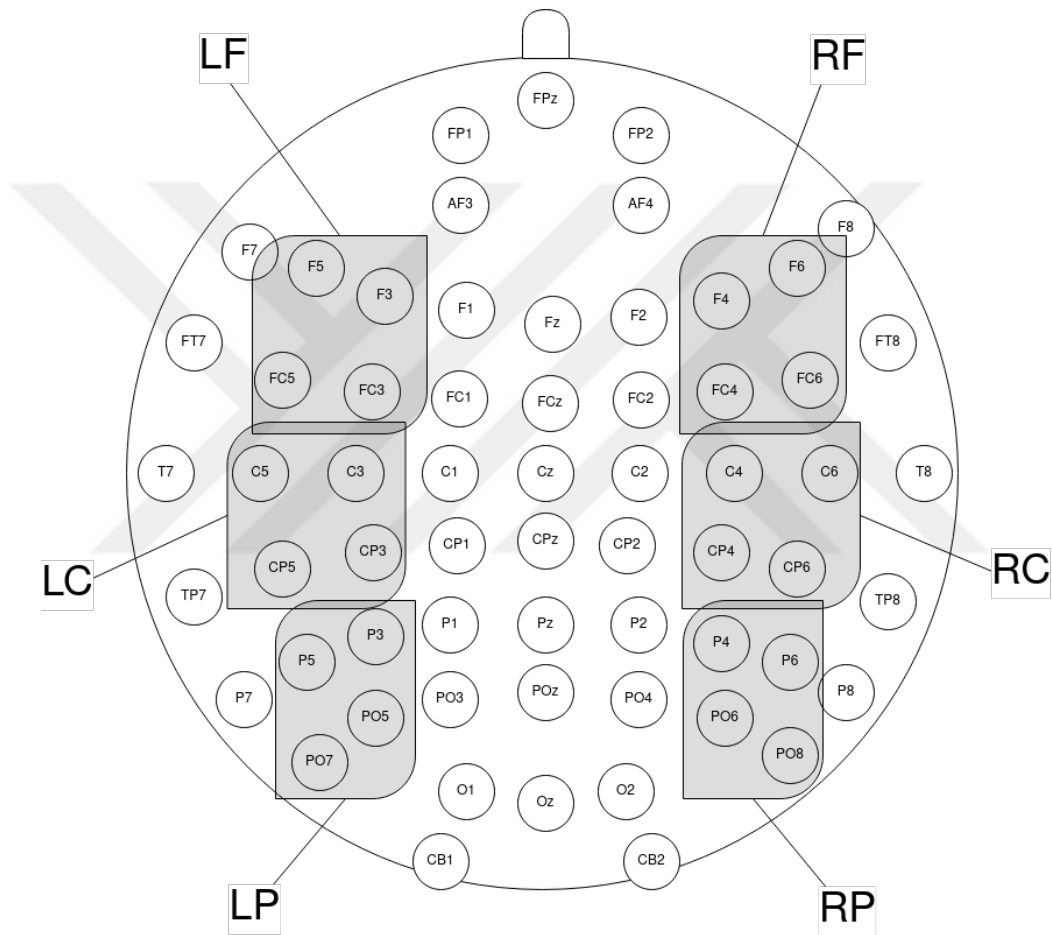


Figure 2.7: Regions of Interest (ROI)

P200 response and a later time window from 450 to 750ms (Laszlo & Federmeier, 2009) was considered after sentence onset for the LPC response. These specific time windows were selected for further analysis, a CSV file was created by taking the average amplitudes of the recorded ERPs for these time windows in each ROI, and statistical analysis was conducted. Figure 2.7 shows the ROIs we have used for the analysis.



Chapter 3

Results

3.1 Statistical Data Analysis

Repeated-measures ANOVAs were used for the ERP analysis. For both the P200 and LPC components, separate 2 (language: L1, L2) x 3 (emotion: neutral, sad, happy) x 2 (accent: native, foreign) x 3 (laterality: right, central, left) x 3 (regions: frontal, central, parietal) analyses were used (Paulmann & Kotz, 2008).

3.2 Results

3.2.1 P200 (170-230ms)

Table 3.1 shows the repeated measure ANOVA results for the P200 component. HeardLang refers to the language of the stimuli, Speaker refers to the accent (Turkish and English speaker) ROI is the Regions of Interest.

Table 3.1: Repeated Measures ANOVA (P200).

Cases	df	F	p	η_p^2
emotion	1.947	0.305	0.732	0.006
emotion; HeardLang	1.947	6.277	0.003	0.108
Residuals	101.219			
ROI	2.320	239.984	<.001	0.822
ROI; HeardLang	2.320	2.454	0.082	0.045
Residuals	120.654			
Speaker; emotion	1.605	0.365	0.649	0.007
Speaker; emotion; HeardLang	1.605	0.697	0.471	0.013
Residuals	83.458			
Speaker; ROI	1.626	8.901	<.001	0.146
Speaker; ROI; HeardLang	1.626	2.087	0.139	0.039
Residuals	84.567			
emotion; ROI	3.673	7.383	<.001	0.124
emotion; ROI; HeardLang	3.673	3.209	0.017	0.058
Residuals	190.988			
Speaker; emotion; ROI	2.483	1.119	0.338	0.021
Speaker; emotion; ROI; HeardLang	2.483	0.437	0.690	0.008
Residuals	129.104			

In the P200 time-window (from 170 to 230 ms) no significant main effect of emotion [$F < 1$] was found. On the other hand, a significant two-way interaction between emotion and ROI [$F(3.673, 190.988) = 7.38, p < 0.01, \eta^2 p = 0.12$] was found.

Post Hoc two-way comparisons (Bonferroni-corrected) at each ROI showed the following results: in LF sites, sentences vocalized with a happy tone showed larger

mean amplitudes compared to the sentences vocalized with both sad and neutral voice (p 's = 0.005). We also observed significant differences between LF happy and LC-LP-RC-RP happy, sad, and neutral (all p 's < .001) such that LF-happy showed larger amplitudes compared to all other conditions and ROIs. There were no significant differences between LF happy and RF-happy, sad and neutral.

There were also significant differences between LF sad and LC-LP-RC-RP happy, sad, and neutral (all p 's < 001). LF-sad showed larger mean amplitudes compared to every emotional condition at LC-LP-RC-RP. However, there were no significant differences between LF sad and RF-happy, sad, and neutral conditions. In addition, LF sad showed descriptively smaller mean amplitudes compared to LF neutral, RF happy, and RF neutral. Larger mean amplitudes were found for LF neutral compared to LC-LP-RC-RP happy, sad, and neutral (all p 's < 001). These differences were not observed for RF-happy, sad and neutral. In addition, LF neutral showed descriptively smaller mean amplitudes compared to RF happy.

In LC sites, there were no significant differences between LC happy and LC-sad and neutral conditions. In addition, there were no significant differences between LC happy and RC sites for any emotional condition. On the other hand, there were significant differences between LC happy and LP-RF-RP happy, sad, and neutral conditions (all p 's < 001). Sentences vocalized with a happy tone of voice in LC sites showed smaller mean amplitudes compared to sentences vocalized with happy, sad, and neutral voice in RF sites, but showed larger mean amplitudes compared to other conditions. Considering LC sites, sentences vocalized with a sad tone of voice descriptively showed smaller mean amplitudes compared to sentences vocalized with a neutral voice, but there was no significant difference between them.

We did find significant differences between LC sad and LP-RF-RP happy, sad, and neutral conditions (all p 's < 001). Within this ROI, sentences vocalized with a sad tone of voice showed larger mean amplitudes compared to happy, sad

and neutral tones of voice from LP and RC sites while showing smaller mean amplitudes compared to happy, sad and neutral voice at RF sites.

We also observed significant differences between LC neutral and LP-RF-RP happy, sad, and neutral (all p 's < 001). There was no significant difference between LC neutral and RC-happy, sad and neutral conditions. LC neutral showed descriptively smaller mean amplitudes compared to all emotional conditions at RF sites.

At LP sites, sentences vocalized with a happy tone of voice showed smaller mean amplitudes compared to sentences vocalized with happy, sad, and neutral voice at the other sites (RF, RC, RP). There was a significant difference between LP happy and LP sad ($p = 0.001$) and a significant difference between RF-RC happy, sad, and neutral (all p 's < 001). On the other hand, at this ROI, sentences vocalized with a happy tone showed larger mean amplitudes compared to RP happy. In addition, there was no significant difference between LP happy and RP-happy, sad, neutral or LP neutral.

For LP sad, there was a marginally significant difference between LP neutral ($p = 0.017$) and a significant difference between LP sad and RF-RC happy, sad, and neutral (all p 's < 001). However, there was no significant difference between LP sad and RP-happy, sad, neutral. In addition, LP sad showed smaller mean amplitudes compared to RF-RC happy, sad and neutral while it showed larger mean amplitudes compared to LP neutral; RP-happy, sad and neutral.

There were also significant differences between LP neutral and RF-RC happy, sad, and neutral (all p 's < 001). However, there was no significant difference between LP neutral and RP-happy, sad, neutral. In addition, LP neutral showed descriptively smaller mean amplitudes compared to RC, RP, RF-happy, sad and neutral except RP happy. In RF sites, sentences vocalized with a happy tone showed larger mean amplitudes compared to sentences vocalized with happy, sad and neutral voices from RC and RP sites (all p 's < 001). On the other hand, there

was no significant difference between RF happy and RF-sad, and neutral ($p = 0.07$; $p = 1$ respectively).

There was a significant difference between RF sad and RC-RP happy, sad, and neutral (all p 's < 0.001). However, there was no significant difference between RF sad and RF neutral ($p = 0.90$). In addition, RF sad showed descriptively larger mean amplitudes compared to RC, RP-happy, sad and neutral except RF neutral. For RF sites, sentences vocalized with neutral tone showed larger mean amplitudes compared to sentences vocalized with happy, sad and neutral voice from RC and RP sites (all p 's < 0.001).

At RC sites, there was also a significant difference found between RC happy and RP happy, sad, and neutral (all p 's < 0.001). On the other hand, there was no significant difference between RC happy and RC-sad and neutral ($p = 1$). RC happy showed descriptively larger mean amplitudes compared to RC sad, and RP-happy, sad, neutral. At this ROI, sentences vocalized with sad and neutral tones of voice showed larger mean amplitudes compared to sentences vocalized with happy, sad and neutral voices from RP sites (all p 's < 0.001). In RP sites, there was a marginally significant difference between RP happy and sad conditions ($p = 0.002$) and it showed smaller mean amplitudes compared to RP sad.

In summary, the interaction between emotion and ROI shows that regardless of emotions, the Left and Right side of the same ROI did not show significant differences.

Additionally, a significant two-way interaction was found between emotion and heard language [$F(1.947, 101.219) = 6.28, p < 0.05, \eta^2 p = 0.11$]. However, Bonferroni-corrected post-hoc tests did not show any significant results for this interaction. The Bonferroni correction is a more conservative method compared to the other post hoc tests. To try and tease apart this interaction, we reverted to using the Tukey correction, but this also did not yield any significant post-hoc results. Based on the descriptive mean differences, we can speculate that

the interaction comes from a difference between the Turkish-happy and Turkish-neutral conditions, such that Turkish-happy has a larger mean amplitude; the Turkish-sad and Turkish-neutral conditions, such that Turkish-sad has a larger mean amplitude; and the English-sad and English-neutral conditions, such that English-neutral has a larger mean amplitude.

A significant main effect regarding regions of interest (ROI) was found [$F(2.320, 120.624) = 239.984, p < 0.01, \eta^2 p = 0.822$]. Post Hoc comparisons (Bonferroni-corrected) at each ROI showed the following results: there were significant differences between LF sites and LC-LP-RC-RP sites (all p 's $< .001$). In addition, there were also significant differences between LC sites and LP-RF-RP (all p 's $< .001$). The LC site showed smaller mean amplitudes compared to RF site. Between LP sites and RF-RC sites we observed significant differences (all p 's $< .001$) LP following a smaller mean amplitude. Finally, RF sites compared to RC-RP sites (all p 's $< .001$) and RC site compared to RP site ($p < .001$) showed larger significant differences.

A marginally significant three-way interaction was found between emotion, ROI and heard language [$F(3.637, 190.988) = 3.21, p = 0.017, \eta^2 p = 0.06$]. Post Hoc three-way comparisons (Bonferroni-corrected) for them showed the following results: at LF sites, English sentences vocalized with happy tone showed larger mean amplitudes compared to English sentences vocalized with sad voice ($p < .001$). In addition, for English-happy-LF, every single LP, LC, RC and RP conditions (combination of each language and emotion) are significantly different, (all p 's $< .001$); English-happy-LF showed larger mean amplitude compared to all other conditions. There was no significant difference between Turkish-happy-LF and other LF conditions. However, Turkish-happy-LF and every single LP, LC, RC and RP conditions (combination of each language and emotion) are significantly different, (all p 's $< .001$), showing larger mean amplitude compared to other conditions. There were marginally significant differences between English-sad-LF compared LC Turkish-sad, Turkish-happy and English-happy conditions

(p 's are .004, .007, and .012 respectively). The English-sad-LF condition showed a larger mean amplitude compared to these conditions. In addition, English-sad-LF and every single LP, RC and RP conditions (combination of each language and emotion) were significantly different, (all p 's < .001). Also, there was a significant difference between English-sad-LF and Turkish neutral-LC (p < .001) and English-sad-LF showed larger mean amplitudes. There were significant differences between Turkish-sad-LF and every single LP, LC, RC and RP conditions (combination of each language and emotion) (all p 's < .001); Turkish-sad-LF showed larger mean amplitude compared to other conditions.

For English and Turkish neutral-LF conditions, significant differences were found compared to every single LP, LC, RC and RP conditions (combination of each language and emotion) (all p 's < .001); and English and Turkish neutral-LF conditions showed larger mean amplitude compared to other conditions.

For English-happy-LC, significant differences were found with every single LP, RF and RP conditions (combination of each language and emotion) (except English-neutral-RF (p = .002) and English-sad-RF (p = .671)) (all other p 's < .001) and English-happy-LC showed larger mean amplitudes compared to each LP and RP conditions.

For Turkish-happy-LC, significant differences were found with every single LP, RF and RP conditions (combination of each language and emotion) (except English-sad-RF (p = .007)) (all p 's < .001) and Turkish-happy-LC showed larger mean amplitudes compared to each LP and RP conditions.

There were significant differences between English-sad-LC and every single LP, RF and RP conditions (combination of each language and emotion) (except English-sad-RF (p = .013)) (all p 's < .001) and English-sad-LC showed larger mean amplitudes compared to LP and RP conditions. However, English-sad-LC showed smaller mean amplitudes compared to each RF conditions. In addition, there were significant differences between Turkish-sad-LC and every single LP, RF

and RP conditions (combination of each language and emotion) (except English-sad-RF ($p = .022$)) (all p 's $< .001$) and Turkish-sad-LC showed larger mean amplitudes compared to LP and RP conditions. However, Turkish-sad-LC showed smaller mean amplitudes compared to each RF conditions.

For English and Turkish neutral-LC conditions, significant differences were found with every single LP, RF and RP conditions (combination of each language and emotion) (except English-sad-RF ($p = .369$) and ($p = .001$) respectively) (all p 's $< .001$); and both English and Turkish neutral-LC conditions showed larger mean amplitude compared to LP and RP conditions. However, both English and Turkish neutral-LC conditions showed smaller mean amplitudes compared to each RF condition. In LP sites, English sentences vocalized with happy tone showed larger mean amplitudes compared to English sentences vocalized with sad voice ($p < .001$). In addition, for English-happy-LP, every single RF and RC conditions (combination of each language and emotion) (all p 's $< .001$); and English-happy-LP showed smaller mean amplitude compared to other conditions.

There were significant differences between Turkish-happy-LP and every single RF and RC conditions (combination of each language and emotion) are significantly different, (all p 's $< .001$); and Turkish-happy-LP showed smaller mean amplitude compared to other conditions. For English and Turkish sad-LP conditions, significant differences were found with every single RF and RC conditions (combination of each language and emotion) (all p 's $< .001$); and Turkish-sad-LP showed smaller mean amplitude compared to other conditions. For English and Turkish neutral-LP conditions, significant differences were found with every single RF and RC conditions (combination of each language and emotion) (all p 's $< .001$); and English and Turkish neutral-LP showed smaller mean amplitude compared to each emotion and language RF and RC conditions.

For English and Turkish happy-RF conditions, significant differences were found with every single RC and RP conditions (combination of each language and

emotion) (all p 's < .001); and English and Turkish happy-RF showed larger mean amplitude compared to other conditions. We found significant differences between English-sad-RF and every single RP condition (combination of each language and emotion), and English RC conditions (all p 's < .001); and English-sad-RF showed larger mean amplitude compared to other conditions. Also, there were marginally significant differences between English-sad-RF and Turkish happy; sad; neutral; RC conditions (p 's = .033; .011; .002 respectively). There were significant differences between Turkish-sad-RF and every single RC, RP conditions (combination of each language and emotion) (all p 's < .001); and Turkish-sad-RF showed larger mean amplitude compared to other conditions.

Both Turkish and English neutral LF were significantly different from every single RC, RP conditions (combination of each language and emotion) are significantly different, (all p 's < .001); and Turkish and English neutral LF showed larger mean amplitude compared to other conditions.

Turkish and English happy RC were significantly different from every single RP condition (combination of each language and emotion) (all p 's < .001); and Turkish and English happy RC showed larger mean amplitude compared to other conditions. Turkish and English sad RC were significantly different from every single RP conditions (combination of each language and emotion) (all p 's < .001); and Turkish and English sad RC showed larger mean amplitude compared to other conditions. Turkish and English neutral RC were significantly different from every single RP conditions (combination of each language and emotion) are significantly different, (all p 's < .001); and Turkish and English neutral RC showed larger mean amplitude compared to other conditions. In RP sites, English sentences vocalized with a happy tone of voice is significantly different from English sentences vocalized with a sad tone of voice (p < .001) and happy vocalized condition showed smaller mean amplitude compared to sad tone of voice condition.

In summary, the interaction between emotion, ROI and heard language shows that regardless of emotions and heard language, the Left and Right side of the same ROI did not show significant differences across conditions. See Figure 3.1 for a grand average of the F3, F4, C3, C4, P3, P4, F5, F6, C5, C6, P5 and P6 waveforms across all conditions for Turkish stimuli and Figure 3.2 for the English stimuli.

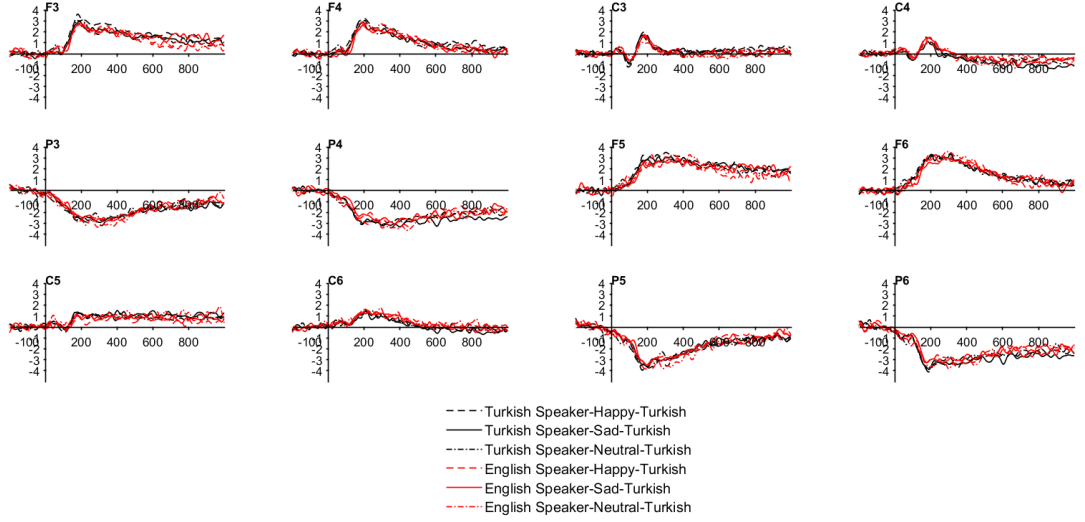


Figure 3.1: Waveforms for Turkish sentences

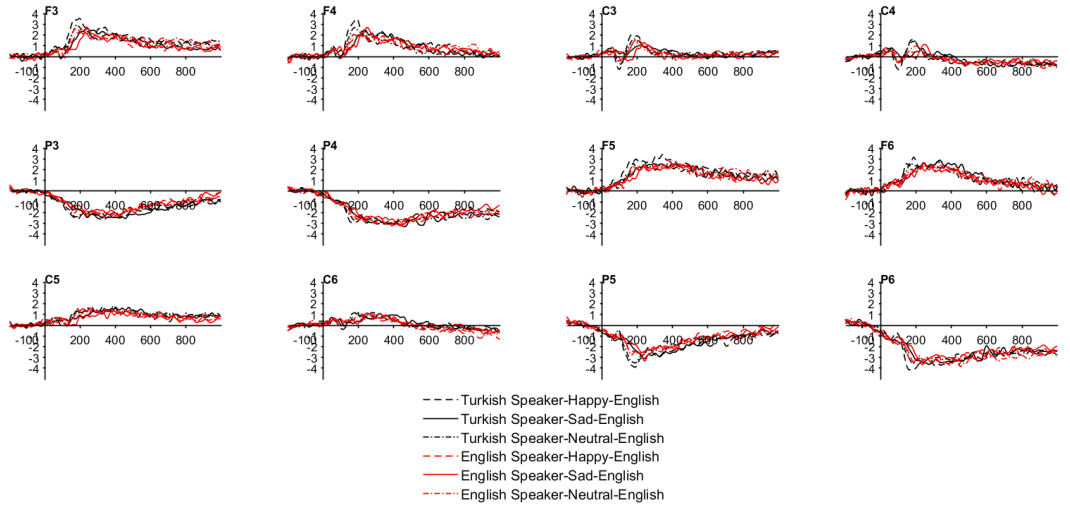


Figure 3.2: Waveforms for English sentences

A significant two-way interaction was found between ROI and speaker [$F(1.626, 84.567) = 8.90, p < 0.01, \eta^2 p = 0.15$]. Post Hoc two-way comparisons (Bonferroni-corrected) at each ROI showed the following results: in LF site, we found a marginally significant difference between the sentences vocalized

by Turkish and English speaker ($p = 0.024$). Turkish speaker LF showed larger mean amplitude compared to English speaker LF. In addition, at this ROI, the significant differences were found between sentences vocalized by Turkish speaker and sentences vocalized by both Turkish and English speaker from LC, LP, RP and RC sites (all p 's $< .001$) and showed larger mean amplitudes.

On the other hand, at this ROI, there were significant differences between sentences vocalized by the English speaker and every LC, LP, RP and RC conditions (both speaker: Turkish and English speaker) (all p 's $< .001$) and showed larger mean amplitudes compared to other conditions. Sentences vocalized by Turkish and English speakers considering the LC sites were significantly different from sentences vocalized by two speakers at LP, RF and RP sites (all p 's $< .001$). They showed smaller mean amplitudes compared to each RF condition while exhibiting larger mean amplitudes for other conditions. At LP sites, sentences vocalized by the Turkish speaker and sentences vocalized by English speaker are marginally significant ($p = .012$) and it showed smaller mean amplitude. In addition, at this ROI, sentences vocalized by both speakers are significantly different compared to sentences vocalized by two speakers at RF and RC sites (all p 's $< .001$). LP conditions showed smaller mean amplitudes compared to RF and RC conditions for either speaker. For RF sites, sentences vocalized by two speakers were significantly different from sentences vocalized by two speakers at RC and RP sites (all p 's $< .001$). RF conditions showed larger mean amplitudes. At RC sites, sentences vocalized by Turkish and English speaker showed larger mean amplitudes compared to sentences vocalized by Turkish and English speaker at RP sites (all p 's $< .001$).

Finally, in RP sites, sentences vocalized by Turkish speaker were marginally significant compared to sentences vocalized by English speaker ($p = .011$) and sentences vocalized by Turkish speaker showed smaller mean amplitude compared to sentences vocalized by English speaker. As a result, the interaction between speaker and ROI shows that different accents did not affect the overall result for

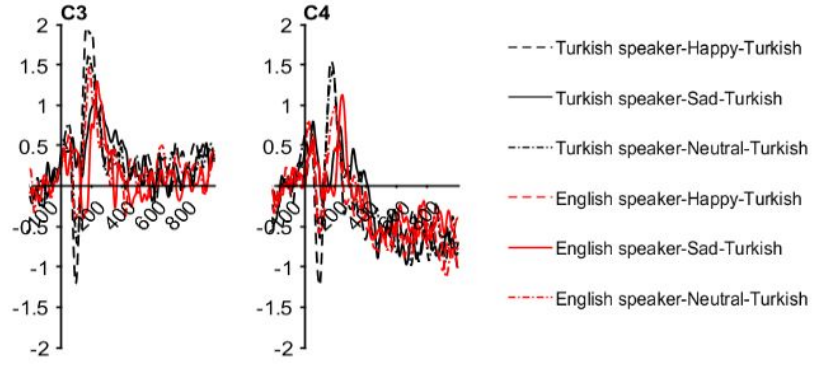


Figure 3.3: Waveforms for Turkish sentences of C3 and C4

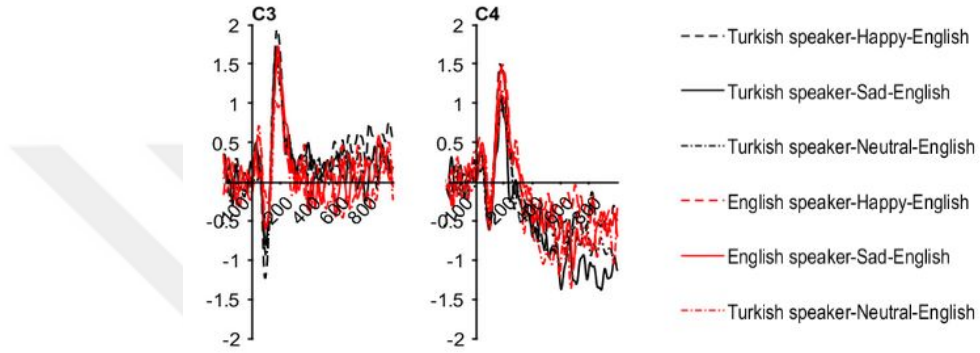


Figure 3.4: Waveforms for English sentences of C3 and C4

this interaction. In addition, the Left and Right side of the same ROI did not show significant differences.

We found a marginally significant two-way interaction between ROI and heard language [$F(2.320, 120.654) = 2.45, p = 0.082, \eta^2 p = 0.05$]. Post Hoc two-way comparisons (Bonferroni-corrected) at each ROI showed the following results: in LF site, we found significant difference between Turkish and English sentences and sentences in both language (Turkish and English) at LC, LP, RP and RC sites (all p 's $< .001$) and showed larger mean amplitudes. In addition, there were significant differences between English and Turkish sentences at LC sites and sentences in both languages at LP, RF, RP sites (all p 's $< .001$). At this ROI, Turkish and English sentences showed larger mean amplitudes compared to every LP and RF conditions while they showed smaller mean amplitudes compared to RF conditions.

For LP sites, sentences in Turkish and English showed smaller mean amplitudes

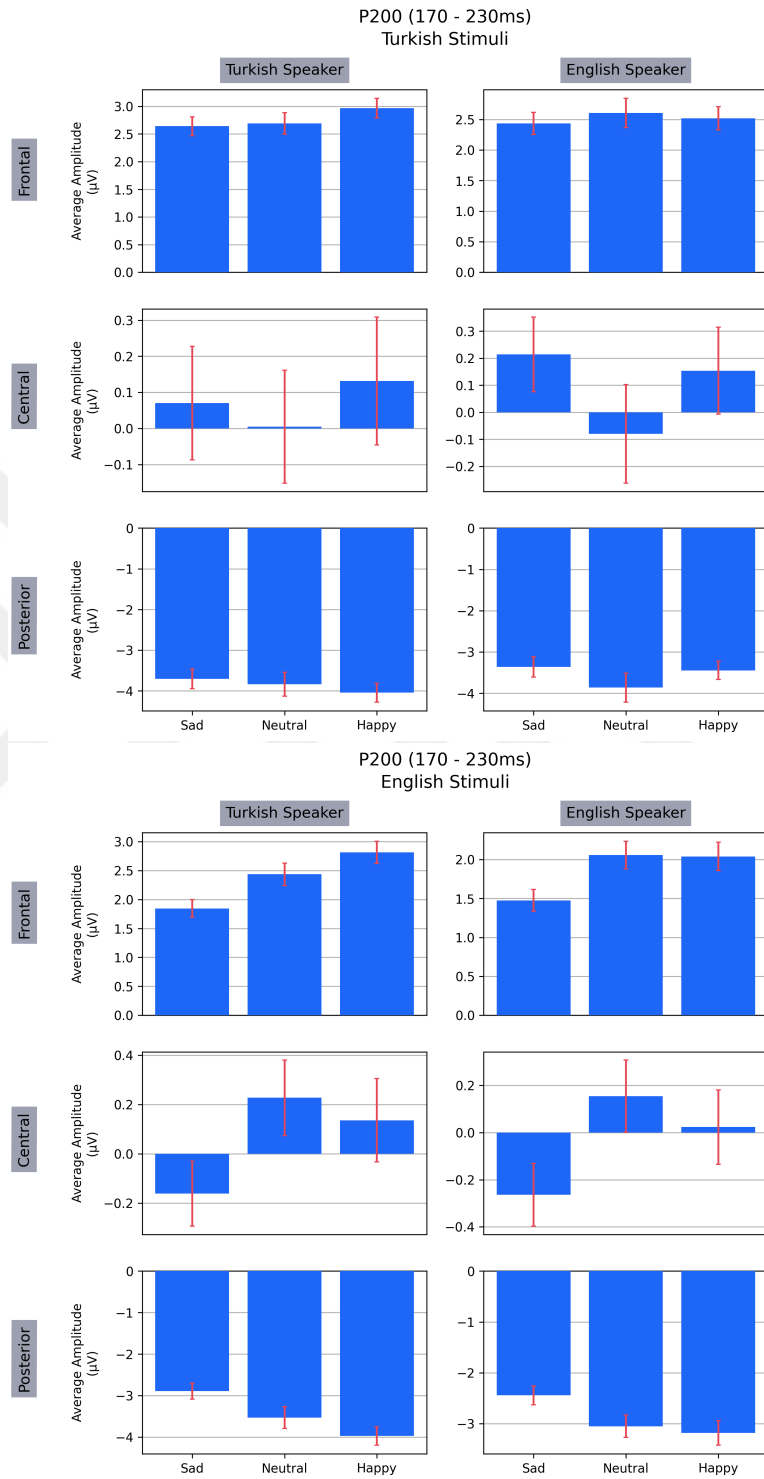


Figure 3.5: P200 mean amplitudes for each emotion at frontal, central and posterior electrode sites. Error bars plot the standard error.

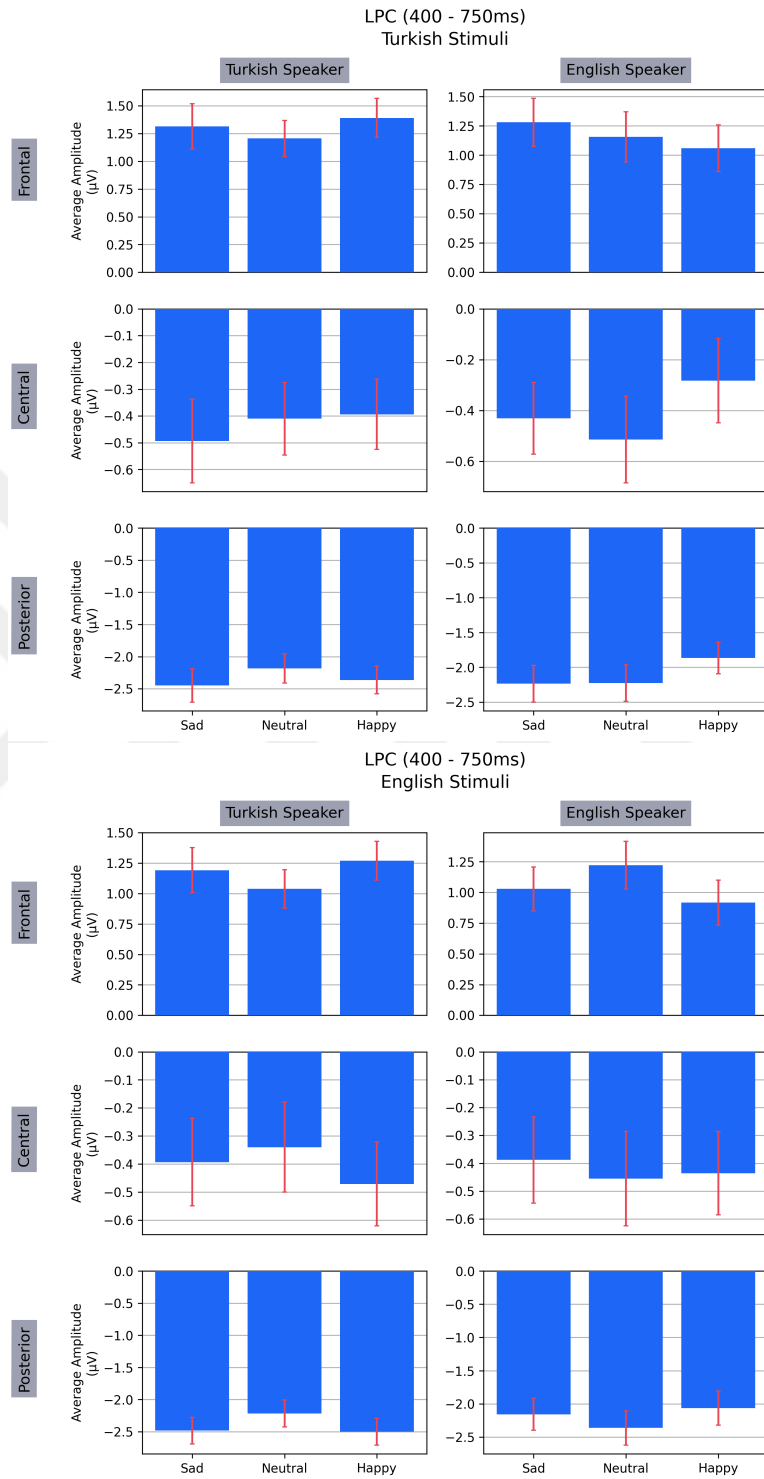


Figure 3.6: LPC mean amplitudes for each emotion at frontal, central and posterior electrode sites. Error bars plot the standard error.

compared to every single condition at RF and RC sites (all p 's $< .001$). Moreover, we observed larger mean amplitudes between at RF sites sentences in Turkish and English and every single condition at RC and RP sites (all p 's $< .001$). Finally, at RC sites, there were significant differences between both Turkish and English sentences and sentences in both languages at RC sites (all p 's $< .001$). See Figure 3.5 and Figure 3.6 for mean P200 and mean LPC amplitudes for each emotion at different electrode sites (frontal, central posterior), across conditions (heard language, accent).

We did not encounter any other significant main effect or interaction during our study. In summary, we confirm through our data analysis that significant differentiation effects are primarily found at frontal and central electrode-sites. Finally, we found no indication that heard language or speaker influences P200 amplitude for our study.

3.2.2 LPC (450-750ms)

In the later time-window (from 450 to 750 ms) no significant main effect of emotion [$F < 1$] was found. However, a significant main effect regarding regions of interest (ROI) was found [$F(1.963, 102.074) = 94.197, p < 0.01, \eta^2 p = 0.644$]. Post Hoc comparisons (Bonferroni-corrected) at each ROI showed the following results: there were significant differences between LF sites and LC-LP-RC-RP sites (all p 's $< .001$) and marginally difference between LF sites and RF sites ($p = .018$). LF sites showed larger mean amplitudes compared to all ROI. In addition, there were also significant differences between LC sites and LP-RP sites (all p 's $< .001$) while marginal difference was found between this ROI and RF-RC sites (p 's are .018 and .003 respectively). LC sites showed smaller mean amplitude compared to RF sites. Moreover, LP sites showed smaller mean amplitudes compared to RF-RC sites (all p 's $< .001$) and marginally difference was found compared to RP sites ($p = .008$). There were significant differences between RF sites and

RC-RP sites (all p 's < 001). Finally, RC sites and RP sites significantly different from each other (p < 001) and it showed larger mean amplitudes.

Table 3.2: Repeated Measures ANOVA (LPC)

Cases	df	F	p	η_p^2
emotion	1.920	0.719	0.484	0.014
emotion; HeardLang	1.920	1.786	0.174	0.033
Residuals	99.851			
ROI	1.963	94.197	<.001	0.644
ROI; HeardLang	1.963	0.219	0.799	0.004
Residuals	102.074			
Speaker; emotion	1.741	1.518	0.226	0.028
Speaker; emotion; HeardLang	1.741	0.201	0.788	0.004
Residuals	90.554			
Speaker; ROI	2.304	2.930	0.050	0.053
Speaker; ROI; HeardLang	2.304	1.677	0.187	0.031
Residuals	119.808			
emotion; ROI	4.163	0.373	0.835	0.007
emotion; ROI; HeardLang	4.163	0.705	0.595	0.013
Residuals	216.473			
Speaker; emotion; ROI	4.223	1.831	0.120	0.034
Speaker; emotion; ROI; HeardLang	4.223	0.212	0.939	0.004
Residuals	219.586			

Table 3.2 shows the repeated measure ANOVA results for the LPC component. HeardLang refers to the language of the stimuli, Speaker refers to the accent (Turkish and English speaker) ROI is the Regions of Interest.

We found a marginally significant two-way interaction between ROI and speaker [$F(2.304, 119.808) = 2.930, p = 0.050, \eta^2 p = 0.053$]. Posthoc two-way comparisons (Bonferroni-corrected) at each ROI showed the following results: in LF site, we found a marginally significant difference between the sentences vocalized by Turkish speaker and sentences vocalized by English speaker at RF sites ($p = 0.018$). Turkish speaker LF showed larger mean amplitude compared to English speaker RF.

In addition, at this ROI, the significant differences were found between sentences vocalized by Turkish speaker and sentences vocalized by both Turkish and

English speaker from LC, LP, RP and RC sites (all p 's < .001) and showed larger mean amplitudes.

On the other hand, at this ROI, there were significant differences between sentences vocalized by English speaker and every LC, LP, RP and RC conditions (both speaker: Turkish and English speaker) (all p 's < .001) and showed larger mean amplitudes compared to other conditions. Sentences vocalized by Turkish and English speaker considering the LC sites were significantly different from sentences vocalized by two speakers at LP and RP sites (all p 's < .001).

They showed smaller mean amplitudes compared to each RF conditions while exhibiting larger mean amplitudes for other conditions. At this ROI, sentences vocalized by both Turkish and English speaker were marginally different from sentences vocalized by Turkish speaker at RC sites (p 's are .007, .018 respectively).

At LP sites, sentences vocalized by Turkish speaker and sentences vocalized by both speakers at RF and RC sites are significantly different (all p 's < .001) and it showed smaller mean amplitudes. In addition, at this ROI, sentences vocalized by English speaker are significantly different compared to sentences vocalized by two speakers at RF sites (all p 's < .001). It showed smaller mean amplitudes compared to RF condition for either speaker. Also, sentences vocalized by English speaker at these sites and every single condition of RC and RP sites are marginally different (p 's are .039, .016, .002, .038 respectively). For RF sites, sentences vocalized by two speakers were significantly different from sentences vocalized by two speakers at RC and RP sites (all p 's < .001). RF conditions showed larger mean amplitudes. Finally, at RC sites, sentences vocalized by Turkish and English speaker showed larger mean amplitudes compared to sentences vocalized by Turkish and English speaker at RP sites (all p 's < .001).

As a result, any other significant main effect or interaction was not found. The interaction between speaker and ROI shows that regardless of accent, the same and across side of the same ROI did not show significant differences.

Chapter 4

Discussion

Prosodic cues, and especially emotional prosody, are an important part of language that help listeners and speakers convey and understand the ultimate meaning of an utterance. According to Wilson and Wharton (2006), emotional prosody is a facilitator for language processing and Schirmer et al. (2005) showed that emotional prosody eases interpretation of sentences and conveys subliminal emotions (i.e. irony, sarcasm) (Paulmann & Pell, 2011). In addition, processing of emotional prosody is socially and biologically important. It might give some important cues to recognize dangerous situations (Öhman & Mineka, 2001) or help us to understand the speech and respond to it appropriately.

In the literature, emotional prosody perception is based on three main hypotheses. These are Universality Hypothesis, In-group Advantage Hypothesis and Dialect Theory. According to universality hypothesis even if a person do not know spoken language and have not been exposed to it before, they can correctly identify conveyed emotions through tone of voice (Scherer et al., 2001). On the other hand, in-group advantage hypothesis refers to people who are good at recognition of vocal emotions when vocalized by in-group members (Van Bezooijen et al., 1983). Finally, Dialect theory addresses the intermediate position, it refers to emotions identified universally however, some expressions have culture

specific aspects (Elfenbein & Ambady, 2003).

This thesis examined the effects of emotional prosody and accents in bilingual participants to answer the following questions:

1. How do bilingual individuals differentially weigh emotional prosody as a cue during speech comprehension in their L1 and L2?
2. How do different accents affect the emotion comprehension of L1 and L2 speakers?
3. Is processing of vocal emotions universal or does it include culturally-specific aspects?

We hypothesized that L2 speakers would use emotional prosody more than L1 speakers because they are not native speakers, and they need more paralinguistic information to arrive at the intended meaning of the speech (Wei et al., 2022; Zhu, 2013). Thus, we would have expected to see larger P200 amplitudes in L2 speakers for emotional stimuli compared to L1 speakers. We also thought that accents can affect emotional prosody processing (based on Floccia et al., 2006; Goslin et al., 2012): participants would easily process native accents in their L1 (a native Turkish speaker speaking Turkish) and they would easily process foreign accents in their L2 (American English speaker speaking English) primarily because our participants are late bilinguals and their exposure has been primarily to accented English. As such, they would be more accustomed to this input in English.

The results of our experiment show that no significant difference was observed between emotional stimuli and neutral stimuli for P200 and LPC components. When we review the literature, the P200 difference is observed between neutral and emotional stimuli in many emotion processing studies conducted. In the literature, P200 is called as an early sign of emotional prosody processing in speech

and this component shows early differentiation between emotional and neutral vocalizations (Schirmer & Kotz, 2006). However, the important point should be considered is higher P200 amplitude was observed when the participant paid attention to emotional stimuli and also higher LPC amplitude was observed during in-depth processing such as meaning assessment. Emotion recognition tasks were assigned to participants in all studies in the literature that had the effect of P200 and LPC (Paulmann & Kotz, 2008; Paulmann et al., 2013; Pell et al., 2009a; Pell & Kotz, 2011). In these studies the forced-choice emotion tasks or the tasks that directly labeled emotions were assigned to participants. This means, the participants were asked to pay attention to the emotional stimuli and to clearly define what the emotion was. However, in our study participants heard semantically neutral sentences that vocalized with different emotional tone of voice and the task did not include any question about emotion recognition. Contrarily, we asked the participants whether the probe word they see on the screen is included in the sentence, which meant we drew their attention to semantics, which may have resulted in completely diverting their attention from paralinguistic cues. Maybe the task and its cognitive demands lead to a situation in which cue usage strategies are more affected than we initially supposed. Thus, presenting vocal stimuli in language that a person has never been exposed to before (e.g. a third unknown language) would allow them to use paralinguistic cues more accurately. However, in this case, we would not be able to examine true linguistic understanding.

In addition, as highlighted above the LPC component is a special component that shows higher amplitude in late time point thanks to the integration of meaning and emotional tones. In this experiment, we could not find the effect in LPC. This might be related to the meaning of sentences and the fact that the emotional vocalization of them was not congruent with their meaning (e.g., there may be a disconnect between hearing a neutral sentence like 'the towel is on the door' in a sad tone). As such, unnatural or anomalous processing conditions may have been created because of the vocalization of very semantically neutral sentences

with different emotional tones of voice.

In the literature, lots of research has found that some basic emotions are recognized better than others (Pell et al., 2009b; Thompson & Balkwill, 2006; Johnson et al., 1986; Scherer et al., 1991). Specifically, Banse and Scherer (1996) showed that sadness and anger are recognized more accurately than others while Zuckerman et al. (1975) and, Paulmann and Pell (2011) found that the least recognized emotion is happiness between languages. On the other hand, other research shows that happiness has the highest recognition rate when it is vocalized by in-group members (Elfenbein & Ambady, 2002). In our experiment, although happy and sad vocalizations were chosen because they are stated as the most salient and least salient in the literature and because they are opposite emotions, we could not find significant difference between emotions. Moreover, our participants did not show processing differences for English and Turkish sentences vocalized by Turkish speaker with a happy tone of voice. This means our findings contradict Elfenbein and Ambady (2002) findings as well.

In addition, our findings contradict with “interlanguage speech intelligibility benefit” phenomenon. This phenomenon examines how individuals’ native language affects the perception of non-native languages. Research shows that people are prone to easily process L1 congruent foreign accented speech in their L2 (Bent & Bradlow, 2003; Hayes-Harb et al., 2008; Imai et al., 2005). According to the interlanguage speech intelligibility benefit phenomenon Turkish–English Bilinguals easily recognize conveyed emotion when English sentences are vocalized by Turkish speakers (L1 congruent foreign accent). However, our results show that participants did not show differences between English sentences vocalized by English speakers or Turkish speakers. These findings are congruent with Stibbard and Lee’s (2006) viewpoint: “foreign accented speech (L1 congruent) in L2 is not always the sign for intelligibility between interlocutors”.

Clarke and Garrett (2004) indicated that English monolingual participants respond slower to English sentences vocalized by Spanish speakers compared to sentences vocalized by English speakers. However, after a one minute exposure to Spanish-accented English stimuli their response rate was getting closer to their response rate to native accented English stimuli. In addition, Strauß et al. (2013) claim that foreign accented speech cause a smaller mean amplitude compared to native accented speech. Furthermore, Romero-Rivas et al. (2015) claim that people are prone to recognizing native accented speech earlier than foreign accented speech, by around 200 ms (P200). However, recognition of foreign accented speech is categorized within itself. For instance, as we move from L1 congruent foreign accent to typologically similar foreign accent, and typologically and prosodically distant foreign accent, the recognition time increases. During our experiments, for the condition when the participants heard Turkish stimuli, participants were repeatedly exposed to Turkish sentences that were vocalized by an American English speaker (foreign Accent). Maybe this over exposure caused a calibration for foreign accented speech and this could be the cause of why we could not observe any difference between native and foreign accented stimuli.

On the other hand, Pell et al. (2009a) claim that basic emotions have similar acoustic features across cultures (e.g. the acoustic properties of fear are similar in different cultures and languages). Although we could not find a significant processing difference in P200 or LPC amplitudes between positive and negative emotions during the experiment, similar acoustic features were found between vocalizations of the same emotion as seen in the acoustic part (Figures 2.2, 2.3 2.4 and Tables 2.3 and 2.4).

In our experiment, when participants heard emotional stimuli vocalized with happy and sad tone in their L1, larger P200 amplitudes were observed compared to stimuli vocalized with neutral tone of voice in our results. However, this situation changed when it came to language processing in participants' L2 they showed larger P200 responses during processing of neutral sentences compared

to emotional sentences. This finding reveals a crucial point: bilingual individuals show heightened emotional processing in their L1. In addition, our results show no interaction between emotion, speaker, and languages. Thus, no difference was found in the participants' processing of Turkish sentences spoken by an American English speaker and Turkish sentences spoken by native Turkish speaker. These results actually draw attention to a very important point in the literature: bilingual individuals differentially process both languages. This point is crucial because it gives insight about bilingual individuals using different strategies for processing and conveying vocal cues in their languages. These strategies might be affected from different factors such as proficiency level and context of language that it used. Bhatara et al. (2016) claim that high L2 proficiency might interfere with emotional prosody processing. As highlighted before, all the participants were recruited from Bilkent University and Middle East Technical university. This means they take their classes in English and their proficiency levels' are at least B2 (upper intermediate) in the Common European Framework of Reference (CEFR). Students of both universities should be accustomed to hearing Turkish-accented English regularly, during their lectures and English-accented Turkish in the campus. This might be the cause of why participants show higher P200 amplitudes for neutral sentences in their L2 instead of emotional sentences. In addition, processing of sentences, especially emotional content in L2, is more difficult than L1 (Dewaele, 2010; Pavlenko, 2005). Since processing sentences in L2 may cause extra load to our participants, it may have been easier to process neutral expressions instead of processing emotional sentences and this might be an explanation for why they show higher P200 amplitudes for neutral stimuli in their L2.

Therefore, all of these findings cause other questions to arise: (1) If these stimuli are vocalized by a foreign speaker of a language the participants have not been exposed to before, would this change the results? (2) Would emotion processing in the L2 change if the participants' L2 was the one to which they were more exposed to emotional content? For question 1, I speculate that if the language

of foreign accent is typologically distant this might cause an extra load to understand speech, so adding this variable to the equation would be helpful to discover the function of typological distant between languages for processing of emotional prosody and accents. Also, for the question 2, I think that exposure to emotional content in one language eases processing of emotional cues because emotions and different features of these emotions are continuously associated in memory and these associations are strengthened across time as claimed in the “three-stage working model for the processing of emotional prosody” (step 3) proposed by Schirmer and Kotz (2006). Therefore, we can speculate that incongruence between our hypothesis and results might be related to the context in which the participant acquired their L2. As I highlighted before, all participants were late bilinguals, and they learned English in school settings. This means they may have limited access to the more emotional aspects of that language. Their experiences are mostly based on the formal side. They are mostly exposed to formal content, and dialogue. This may explain why the participants showed a preference toward neutral tone of voice in their L2, regardless of accent.

Another reason might be the availability of more than one paralinguistic cue might interfere with each other. In reality, a lot of information is embedded in the vocal cues. For instance, when listeners hear the speech without any visualization, they can easily get lots of information such as gender, approximate age, emotional state, even socio-economic status (mostly thanks to accent) of speakers (Belin et al., 2004). Research shows that emotional prosody and accent are two crucial vocal cues (Spreckelmeyer et al., 2009). So, during the processing of these vocal cues might interfere with each other and cause delays in processing this information.

4.1 Limitations and Future Direction

One main limitation of this research was the number of presented accents and emotions. I believe that if more emotional vocalization stimuli had been presented to the participants, the results would have been more generalizable to other groups of participants. In the literature, most vocal emotional processing studies are conducted with few emotions and these emotions are not balanced (not equal positive and negative emotions). Although our research was conducted with balanced emotions (one neutral, one positive, and one negative emotion), examining more emotions might give insight into how different emotions are processed differently in bilinguals' respective L1 and L2, especially if these stimuli are vocalized with different accents. According to the in-group advantages hypothesis, positive emotions are recognized better when they are vocalized by in-group members. Therefore, seeing the effect of different emotions being vocalized by both in-group and out-group members on emotion processing and understanding the saliency differences between different emotions would provide important results for the vocal emotion literature. In future studies, it may be useful to recruit a native speaker from a language that the participants have never been exposed to before and use their speech as foreign accented speech.

Another limitation is that all participants were late bilinguals. I think that a comparison between different types of bilinguals would be very important for understanding the nature of emotion processing. So recruiting different types of bilinguals would be beneficial for future studies.

On the other hand, the structure of our sentences would create a problem during sentence comprehension. In terms of linguistic analyses of sentences, some stimuli were passive sentences that create an extra load, especially for L2 speakers because of complexity of sentences i.e. The dress hung on a hanger. Furthermore, our sentences do not have an identical beginning, they started with different

complex words (pronouns, nouns, etc.). This situation could cause comprehension and processing differences, especially for studies conducted in different languages. To avoid these linguistic confoundings more homogeneous sentences and similar structures should be used in future studies.

On the methodological side, creating completely semantically neutral sentences and vocalizing these sentences with an emotional tone of voice causes a confounding variable. Just as vocalizing a happy sentence with a sad tone causes an incongruency between what the listener understands semantically and what they feel emotionally, it may be the case that vocalizing semantically neutral sentences emotionally causes the same problem. For example, why would a person need to vocalize “The green towel is on the table” in a sad or happy tone? In order to clarify this problem, I think it would be useful to have a pseudoword condition such that there is no semantics attached to this condition; comparing this with a semantically congruent condition would clarify this potential confounding. Another alternative would be the usage of sentences that would be more appropriate to produce emotionally for these kinds of experiments. For instance, “Alice came” might be a great example. The arrival of Alice might be a good or bad thing so vocalization with different emotions would not create incongruency between meaning and emotional vocalization.

On the other hand, two different norming studies were conducted for this research to be sure about whether the sentences are semantically neutral and whether the vocalization of sentences conveyed the target emotions (emotions that we expect i.e. happily vocalized sentences heard as happy sentences). Although we normed the sentences with native Turkish speakers, we cannot exclude the possibility that the listeners in the EEG experiment perceived them differently and did not ‘recognize’ the emotions in the same way as those that were in Norming Study 2. Thus, for future studies assigning an emotional task would be crucial to be sure whether the participants correctly recognized and processed conveyed emotions during the EEG session

4.2 Conclusion

Consequently, our results did not support our hypothesis about the usage of vocal emotional cues. However, the hypothesis about the contents of language was supported, and we saw that L1 and L2 speakers differently process emotional prosody in their L1 and L2. In addition, our findings support the in-group advantage, the nativelikeness of stimuli elicit higher P200 amplitudes for emotionally vocalized sentences. However, the accent is not a game changer, it does not affect emotional prosody perception. This study also revealed another important aspect for EEG studies: if the task given to the participants does not include overt emotion recognition, vocal emotional cues may not provide enough information to elicit a strong P200 amplitude. Therefore, our findings trigger lots of important questions for future studies and shed light on emotional prosody processing theories and perception of different vocal cues.

References

- Albas, D. C. et al. (1976). "Perception of the Emotional Content of Speech: A Comparison of Two Canadian Groups". In: *Journal of Cross-Cultural Psychology* 7, pp. 481–490. ISSN: 1552-5422. DOI: 10.1177/002202217674009.
- Altarriba, J. and Basnight-Brown, D. M. (Sept. 2011). "The Representation of Emotion vs. Emotion-Laden Words in English and Spanish in the Affective Simon Task". In: *International Journal of Bilingualism* 15.3, pp. 310–328. ISSN: 1367-0069. DOI: 10.1177/1367006910379261. (Visited on 05/12/2023).
- Altmann, U. et al. (2012). "The Power of Emotional Valence-from Cognitive to Affective Processes in Reading". In: *Frontiers in Human Neuroscience* 6, p. 192. ISSN: 1662-5161. DOI: 10.3389/fnhum.2012.00192.
- Altrov, R (2013). "Aspects of Cultural Communication in Recognizing Emotions". In: *Trames. Journal of the Humanities and Social Sciences* 17.2, p. 159. ISSN: 1406-0922. DOI: 10.3176/tr.2013.2.04. (Visited on 05/12/2023).
- Anooshian, L. J. and Hertel, P. T. (Nov. 1994). "Emotionality in Free Recall: Language Specificity in Bilingual Memory". In: *Cognition and Emotion* 8.6, pp. 503–514. ISSN: 0269-9931. DOI: 10.1080/02699939408408956. (Visited on 05/12/2023).
- Ashley, V. et al. (Jan. 2004). "Time Course and Specificity of Event-Related Potentials to Emotional Expressions". In: *NeuroReport* 15.1, p. 211. ISSN: 0959-4965. (Visited on 05/12/2023).

- Ayçiçeği-Dinn, A. and Caldwell-Harris, C. L. (July 2009). “Emotion-Memory Effects in Bilingual Speakers: A Levels-of-Processing Approach”. In: *Bilingualism: Language and Cognition* 12.3, pp. 291–303. ISSN: 1469-1841, 1366-7289. DOI: 10.1017/S1366728909990125. (Visited on 05/12/2023).
- Ayçiçeği, A. and Harris, C. L. (2004). “Brief Report: Bilinguals’ Recall and Recognition of Emotion Words”. In: *Cognition and Emotion* 18, pp. 977–987. ISSN: 1464-0600. DOI: 10.1080/02699930341000301.
- Banise, R. and Scherer, K. R. (1996). “Acoustic Profiles in Vocal Emotion Expression”. In: *Journal of Personality and Social Psychology* 70, pp. 614–636. ISSN: 1939-1315. DOI: 10.1037/0022-3514.70.3.614.
- Beach, C. M. (Dec. 1991). “The Interpretation of Prosodic Patterns at Points of Syntactic Structure Ambiguity: Evidence for Cue Trading Relations”. In: *Journal of Memory and Language* 30.6, pp. 644–663. ISSN: 0749-596X. DOI: 10.1016/0749-596X(91)90030-N. (Visited on 05/12/2023).
- Beier, E. G. and Zautra, A. J. (1972). “Identification of Vocal Communication of Emotions across Cultures”. In: *Journal of Consulting and Clinical Psychology* 39, pp. 166–166. ISSN: 1939-2117. DOI: 10.1037/h0033170.
- Belin, P. et al. (Mar. 2004). “Thinking the Voice: Neural Correlates of Voice Perception”. In: *Trends in Cognitive Sciences* 8.3, pp. 129–135. ISSN: 1364-6613. DOI: 10.1016/j.tics.2004.01.008. (Visited on 05/12/2023).
- Ben-Zeev, S. (1977). “The Influence of Bilingualism on Cognitive Strategy and Cognitive Development”. In: *Child Development* 48.3, pp. 1009–1018. ISSN: 0009-3920. DOI: 10.2307/1128353. JSTOR: 1128353. (Visited on 05/12/2023).
- Bent, T. and Bradlow, A. R. (Aug. 2003). “The Interlanguage Speech Intelligibility Benefit”. In: *The Journal of the Acoustical Society of America* 114.3, pp. 1600–1610. ISSN: 0001-4966. DOI: 10.1121/1.1603234. (Visited on 05/12/2023).
- Berckmoes, C. and Vingerhoets, G. (Oct. 2004). “Neural Foundations of Emotional Speech Processing”. In: *Current Directions in Psychological Science* 13.5,

- pp. 182–185. ISSN: 0963-7214. DOI: 10.1111/j.0963-7214.2004.00303.x. (Visited on 05/12/2023).
- Besson, M. et al. (Oct. 2002). “Emotional Prosody: Sex Differences in Sensitivity to Speech Melody”. In: *Trends in Cognitive Sciences* 6.10, pp. 405–407. ISSN: 1879-307X. DOI: 10.1016/s1364-6613(02)01975-7.
- Bhatara, A. et al. (2016). “Second Language Ability and Emotional Prosody Perception”. In: *PLoS ONE* 11. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0156855.
- Boersma, P. and van Heuven, V. (2001). “Speak and unSpeak with PRAAT”. In: 5.9.
- Bostanov, V. and Kotchoubey, B. (2004). “Recognition of Affective Prosody: Continuous Wavelet Measures of Event-Related Brain Potentials to Emotional Exclamations”. In: *Psychophysiology* 41.2, pp. 259–268. ISSN: 1469-8986. DOI: 10.1111/j.1469-8986.2003.00142.x. (Visited on 05/12/2023).
- Bower, G. H. (1981). “Mood and Memory”. In: *American Psychologist* 36, pp. 129–148. ISSN: 1935-990X. DOI: 10.1037/0003-066X.36.2.129.
- (1987). “Commentary on Mood and Memory”. In: *Behaviour Research and Therapy* 25, pp. 443–455. ISSN: 1873-622X. DOI: 10.1016/0005-7967(87)90052-0.
- Bruce, V. and Young, A. (1986). “Understanding Face Recognition”. In: *British Journal of Psychology* 77, pp. 305–327. ISSN: 2044-8295. DOI: 10.1111/j.2044-8295.1986.tb02199.x.
- Bryant, G. A. and Barrett, H. C. (2008). “Vocal Emotion Recognition across Disparate Cultures”. In: *Journal of Cognition and Culture* 8, pp. 135–148. ISSN: 1568-5373. DOI: 10.1163/156770908X289242.
- Chen, X. et al. (Mar. 2011). “Event-Related Potential Correlates of the Expectancy Violation Effect during Emotional Prosody Processing”. In: *Biological Psychology* 86.3, pp. 158–167. ISSN: 0301-0511. DOI: 10.1016/j.biopsycho.2010.11.004. (Visited on 05/12/2023).

- Clahsen, H. and Felser, C. (Jan. 2006). “Continuity and Shallow Structures in Language Processing”. In: *Applied Psycholinguistics* 27.1, pp. 107–126. ISSN: 1469-1817, 0142-7164. DOI: 10.1017/S0142716406060206. (Visited on 05/12/2023).
- Clarke, C. M. and Garrett, M. F. (Dec. 2004). “Rapid Adaptation to Foreign-Accented English”. In: *The Journal of the Acoustical Society of America* 116.6, pp. 3647–3658. ISSN: 0001-4966. DOI: 10.1121/1.1815131. (Visited on 05/12/2023).
- Comeau, L. et al. (June 2003). “The Modeling Hypothesis and Child Bilingual Codemixing”. In: *International Journal of Bilingualism* 7.2, pp. 113–126. ISSN: 1367-0069. DOI: 10.1177/13670069030070020101. (Visited on 05/12/2023).
- Comeau, L. et al. (2007). “Bilingual Children’s Repairs of Breakdowns in Communication”. In: *Journal of Child Language* 34, pp. 159–174. ISSN: 1469-7602. DOI: 10.1017/S0305000906007690.
- Conrad, M. et al. (2011). “The Time Course of Emotion Effects in First and Second Language Processing: A Cross Cultural ERP Study with German–Spanish Bilinguals”. In: *Frontiers in Psychology* 2. ISSN: 1664-1078. (Visited on 05/12/2023).
- Contemori, C. and Tortajada, F. (Jan. 2020). “The Use of Social–Communicative Cues to Interpret Ambiguous Pronouns: Bilingual Adults Differ from Monolinguals”. In: *Applied Psycholinguistics* 41.1, pp. 51–77. ISSN: 0142-7164, 1469-1817. DOI: 10.1017/S0142716419000407. (Visited on 05/12/2023).
- Cordaro, D. T. et al. (2016). “The Voice Conveys Emotion in Ten Globalized Cultures and One Remote Village in Bhutan”. In: *Emotion* 16, pp. 117–128. ISSN: 1931-1516. DOI: 10.1037/emo0000100.
- Cowie, R. and Cornelius, R. R. (2003). “Describing the Emotional States That Are Expressed in Speech”. In: *Speech Communication* 40, pp. 5–32. ISSN: 1872-7182. DOI: 10.1016/S0167-6393(02)00071-7.

- Cristia, A. et al. (2012). "Linguistic Processing of Accented Speech Across the Lifespan". In: *Frontiers in Psychology* 3. ISSN: 1664-1078. (Visited on 05/12/2023).
- Cummins, J. and Mulcahy, R. (1978). "Orientation to Language in Ukrainian-English Bilingual Children". In: *Child Development* 49, pp. 1239–1242. ISSN: 1467-8624. DOI: 10.2307/1128769.
- Cummings, I. (2017). "Parsing and Working Memory in Bilingual Sentence Processing". In: *Bilingualism: Language and Cognition* 20, pp. 659–678. ISSN: 1469-1841. DOI: 10.1017/S1366728916000675.
- Degner, J. et al. (Jan. 2012). "It Matters How Much You Talk: On the Automaticity of Affective Connotations of First and Second Language Words". In: *Bilingualism: Language and Cognition* 15.1, pp. 181–189. ISSN: 1469-1841, 1366-7289. DOI: 10.1017/S1366728911000095. (Visited on 05/12/2023).
- Deniz, N. D. (Nov. 2022). "Processing Syntactic and Semantic Information in the L2: Evidence for Differential Cue-Weighting in the L1 and L2". In: *Bilingualism: Language and Cognition* 25.5, pp. 713–725. ISSN: 1366-7289, 1469-1841. DOI: 10.1017/S1366728921001140. (Visited on 05/12/2023).
- Dewaele, J.-M. (2004). "Perceived Language Dominance and Language Preference for Emotional Speech: The Implications for Attrition Research". In: *Studies in Bilingualism*. Ed. by M. S. Schmid et al. Vol. 28. Amsterdam: John Benjamins Publishing Company, p. 81. ISBN: 978-90-272-4139-9 978-1-58811-549-2 978-90-272-9527-9. DOI: 10.1075/sibil.28.06dew. (Visited on 05/12/2023).
- (2010). *Emotions in Multiple Languages*. London: Palgrave Macmillan UK. ISBN: 978-1-349-52184-5 978-0-230-28950-5. DOI: 10.1057/9780230289505. (Visited on 05/12/2023).
- Dromey, C. et al. (Nov. 2005). "Recognition of Affective Prosody by Speakers of English as a First or Foreign Language". In: *Speech Communication* 47.3, pp. 351–359. ISSN: 0167-6393. DOI: 10.1016/j.specom.2004.09.010. (Visited on 05/12/2023).

- Eilola, T. M. et al. (Aug. 2007). “Emotional Activation in the First and Second Language”. In: *Cognition and Emotion* 21.5, pp. 1064–1076. ISSN: 0269-9931. DOI: 10.1080/02699930601054109. (Visited on 05/12/2023).
- Ekman, P. et al. (Apr. 1969). “Pan-Cultural Elements in Facial Displays of Emotion”. In: *Science (New York, N.Y.)* 164.3875, pp. 86–88. ISSN: 0036-8075. DOI: 10.1126/science.164.3875.86.
- Ekman, P. (1992). “An Argument for Basic Emotions”. In: *Cognition and Emotion* 6, pp. 169–200. ISSN: 1464-0600. DOI: 10.1080/02699939208411068.
- (1993). “Facial Expression and Emotion”. In: *American Psychologist* 48, pp. 384–392. ISSN: 1935-990X. DOI: 10.1037/0003-066X.48.4.384.
- (1994). “Strong Evidence for Universals in Facial Expressions: A Reply to Russell’s Mistaken Critique”. In: *Psychological Bulletin* 115, pp. 268–287. ISSN: 1939-1455. DOI: 10.1037/0033-2909.115.2.268.
- Ekman, P. and Friesen, W. V. (1971). “Constants across Cultures in the Face and Emotion”. In: *Journal of Personality and Social Psychology* 17, pp. 124–129. ISSN: 1939-1315. DOI: 10.1037/h0030377.
- Elfenbein, H. A. and Ambady, N. (Mar. 2002). “On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis”. In: *Psychological Bulletin* 128.2, pp. 203–235. ISSN: 0033-2909. DOI: 10.1037/0033-2909.128.2.203.
- (2003). “Universals and Cultural Differences in Recognizing Emotions of a Different Cultural Group”. In: *Current Directions in Psychological Science* 12, pp. 159–164. ISSN: 1467-8721. DOI: 10.1111/1467-8721.01252.
- Elfenbein, H. A. et al. (2007). “Toward a Dialect Theory: Cultural Differences in the Expression and Recognition of Posed Facial Expressions”. In: *Emotion* 7, pp. 131–146. ISSN: 1931-1516. DOI: 10.1037/1528-3542.7.1.131.
- Fairbanks, G. and Pronovost, W. (Oct. 1938). “Vocal Pitch During Simulated Emotion”. In: *Science* 88.2286, pp. 382–383. DOI: 10.1126/science.88.2286.382. (Visited on 05/12/2023).

- Ferré, P. et al. (July 2013). “Memory for Emotional Words in the First and the Second Language: Effects of the Encoding Task”. In: *Bilingualism: Language and Cognition* 16.3, pp. 495–507. ISSN: 1366-7289, 1469-1841. DOI: 10.1017/S1366728912000314. (Visited on 05/12/2023).
- Ferré, P. et al. (Nov. 2022). “Love Me in L1, but Hate Me in L2: How Native Speakers and Bilinguals Rate the Affectivity of Words When Feeling or Thinking about Them”. In: *Bilingualism: Language and Cognition* 25.5, pp. 786–800. ISSN: 1366-7289, 1469-1841. DOI: 10.1017/S1366728922000189. (Visited on 05/12/2023).
- Fiske, A. P. et al. (1998). “The Cultural Matrix of Social Psychology”. In: *The Handbook of Social Psychology, Vols. 1-2, 4th Ed.* New York, NY, US: McGraw-Hill, pp. 915–981. ISBN: 978-0-19-521376-8.
- Flege, J. E. (1995). “Flege_in_Strange_1995.Pdf”. In: *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. York Press.
- Floccia, C. et al. (2006). “Does a Regional Accent Perturb Speech Processing?” In: *Journal of Experimental Psychology: Human Perception and Performance* 32, pp. 1276–1293. ISSN: 1939-1277. DOI: 10.1037/0096-1523.32.5.1276.
- Freeman, M. R. et al. (2016). “Cognitive and Emotional Effects of Bilingualism in Adulthood”. In: *Bilingualism across the Lifespan: Factors Moderating Language Proficiency*. Language and the Human Lifespan Series. Washington, DC, US: American Psychological Association, pp. 285–303. ISBN: 978-1-4338-2283-4 978-1-4338-2284-1. DOI: 10.1037/14939-016.
- Fridlund, A. J. (1994). *Human Facial Expression: An Evolutionary View*. Human Facial Expression: An Evolutionary View. San Diego, CA, US: Academic Press, pp. xiv, 369. ISBN: 978-0-12-267630-7.
- Fuertes, J. N. et al. (2012). “A Meta-Analysis of the Effects of Speakers’ Accents on Interpersonal Evaluations”. In: *European Journal of Social Psychology* 42.1, pp. 120–133. ISSN: 1099-0992. DOI: 10.1002/ejsp.862. (Visited on 05/12/2023).

- Genesee, F. et al. (1975). "Communication Skills of Bilingual Children". In: *Child Development* 46.4, pp. 1010–1014. ISSN: 0009-3920. DOI: 10.2307/1128415. JSTOR: 1128415. (Visited on 05/12/2023).
- Giles, H. and Billings, A. C. (2004). "Assessing Language Attitudes: Speaker Evaluation Studies". In: *The handbook of applied linguistics*, pp. 187–209.
- Gluszek, A. and Dovidio, J. F. (May 2010). "The Way They Speak: A Social Psychological Perspective on the Stigma of Nonnative Accents in Communication". In: *Personality and Social Psychology Review* 14.2, pp. 214–237. ISSN: 1088-8683, 1532-7957. DOI: 10.1177/1088868309359288. (Visited on 05/12/2023).
- Goslin, J. et al. (Aug. 2012). "An ERP Investigation of Regional and Foreign Accent Processing". In: *Brain and Language* 122.2, pp. 92–102. ISSN: 0093-934X. DOI: 10.1016/j.bandl.2012.04.017. (Visited on 05/13/2023).
- Göz, İ. et al. (Oct. 2017). "Subjective Age-of-Acquisition Norms for 600 Turkish Words from Four Age Groups". In: *Behavior Research Methods* 49.5, pp. 1736–1746. ISSN: 1554-3528. DOI: 10.3758/s13428-016-0817-y. (Visited on 05/12/2023).
- Graham, C. R. et al. (Apr. 2001). "Recognition of Emotion in English Voices by Speakers of Japanese, Spanish and English". In: 39.1, pp. 19–37. ISSN: 1613-4141. DOI: 10.1515/iral.39.1.19. (Visited on 05/12/2023).
- Grandjean, D. (Jan. 2021). "Brain Networks of Emotional Prosody Processing". In: *Emotion Review* 13.1, pp. 34–43. ISSN: 1754-0739. DOI: 10.1177/1754073919898522. (Visited on 05/12/2023).
- Grey, S. et al. (Nov. 2019). "Processing Foreign-Accented Speech in a Second Language: Evidence from ERPs during Sentence Comprehension in Bilinguals". In: *Bilingualism: Language and Cognition* 22.5, pp. 912–929. ISSN: 1366-7289, 1469-1841. DOI: 10.1017/S1366728918000937. (Visited on 05/12/2023).
- Grüter, T. et al. (Feb. 2020). "How Classifiers Facilitate Predictive Processing in L1 and L2 Chinese: The Role of Semantic and Grammatical Cues". In: *Language, Cognition and Neuroscience* 35.2, pp. 221–234. ISSN: 2327-3798. DOI: 10.1080/23273798.2019.1648840. (Visited on 05/12/2023).

- Hatzidaki, A. et al. (2015). "The Way You Say It, the Way I Feel It: Emotional Word Processing in Accented Speech". In: *Frontiers in Psychology* 6. ISSN: 1664-1078. (Visited on 05/12/2023).
- Hayes-Harb, R. et al. (Oct. 2008). "The Interlanguage Speech Intelligibility Benefit for Native Speakers of Mandarin: Production and Perception of English Word-Final Voicing Contrasts". In: *Journal of Phonetics* 36.4, pp. 664–679. ISSN: 0095-4470. DOI: 10.1016/j.wocn.2008.04.002. (Visited on 05/12/2023).
- Heyrani, R. et al. (2021). "Laterality in Emotional Language Processing in First and Second Language". In: *Frontiers in Psychology* 12, p. 736359. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2021.736359.
- Imai, S. et al. (Jan. 2005). "Lexical Frequency and Neighborhood Density Effects on the Recognition of Native and Spanish-accented Words by Native English and Spanish Listeners". In: *The Journal of the Acoustical Society of America* 117.2, pp. 896–907. ISSN: 0001-4966. DOI: 10.1121/1.1823291. (Visited on 05/12/2023).
- Izard, C. E. (1994). "Innate and Universal Facial Expressions: Evidence from Developmental and Cross-Cultural Research". In: *Psychological Bulletin* 115, pp. 288–299. ISSN: 1939-1455. DOI: 10.1037/0033-2909.115.2.288.
- Johnson, W. F. et al. (1986). "Recognition of Emotion from Vocal Cues". In: *Archives of General Psychiatry* 43, pp. 280–283. ISSN: 1538-3636. DOI: 10.1001/archpsyc.1986.01800030098011.
- Juslin, P. N. and Laukka, P. (2003). "Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?" In: *Psychological Bulletin* 129, pp. 770–814. ISSN: 1939-1455. DOI: 10.1037/0033-2909.129.5.770.
- Kagan, J. (Dec. 2007). "A Trio of Concerns". In: *Perspectives on Psychological Science* 2.4, pp. 361–376. ISSN: 1745-6916, 1745-6924. DOI: 10.1111/j.1745-6916.2007.00049.x. (Visited on 05/12/2023).

- Kanske, P. and Kotz, S. A. (May 2007). “Concreteness in Emotional Words: ERP Evidence from a Hemifield Study”. In: *Brain Research* 1148, pp. 138–148. ISSN: 0006-8993. DOI: 10.1016/j.brainres.2007.02.044. (Visited on 05/12/2023).
- Kensinger, E. A. and Corkin, S. (2003). “Memory Enhancement for Emotional Words: Are Emotional Words More Vividly Remembered than Neutral Words?” In: *Memory & Cognition* 31, pp. 1169–1180. ISSN: 1532-5946. DOI: 10.3758/BF03195800.
- Kinzler, K. D. et al. (Aug. 2009). “Accent Trumps Race in Guiding Children’s Social Preferences”. In: *Social Cognition* 27.4, pp. 623–634. ISSN: 0278-016X. DOI: 10.1521/soco.2009.27.4.623. (Visited on 05/12/2023).
- Kramer, E. (1964). “Elimination of Verbal Cues in Judgments of Emotion from Voice”. In: *The Journal of Abnormal and Social Psychology* 68, pp. 390–396. ISSN: 0096-851X. DOI: 10.1037/h0042473.
- Laszlo, S. and Federmeier, K. D. (2009). “A Beautiful Day in the Neighborhood: An Event-Related Potential Study of Lexical Relationships and Prediction in Context”. In: *Journal of Memory and Language* 61.3, pp. 326–338.
- Lev-Ari, S. and Keysar, B. (Nov. 2010). “Why Don’t We Believe Non-Native Speakers? The Influence of Accent on Credibility”. In: *Journal of Experimental Social Psychology* 46.6, pp. 1093–1096. ISSN: 0022-1031. DOI: 10.1016/j.jesp.2010.05.025. (Visited on 05/12/2023).
- Levitt, E. A. (1964). “The Relationship between Abilities to Express Emotional Meanings Vocally and Facially”. In: *The communication of emotional meaning. New York: McGraw-Hill* 19, p. 64.
- Li, P. et al. (Nov. 2020). “Language History Questionnaire (LHQ3): An Enhanced Tool for Assessing Multilingual Experience”. In: *Bilingualism: Language and Cognition* 23.5, pp. 938–944. ISSN: 1366-7289, 1469-1841. DOI: 10.1017/S1366728918001153. (Visited on 05/12/2023).
- Lindquist, K. A. et al. (2015). “Does Language Do More than Communicate Emotion?” In: *Current Directions in Psychological Science* 24, pp. 99–108. ISSN: 1467-8721. DOI: 10.1177/0963721414553440.

- Liu, T. et al. (Jan. 2012). “Electrophysiological Insights into Processing Nonverbal Emotional Vocalizations”. In: *NeuroReport* 23.2, p. 108. ISSN: 0959-4965. DOI: 10.1097/WNR.0b013e32834ea757. (Visited on 05/12/2023).
- Luk, G. et al. (2011). “Is There a Relation between Onset Age of Bilingualism and Enhancement of Cognitive Control?” In: *Bilingualism: Language and Cognition* 14, pp. 488–495.
- Major, R. C. et al. (2002). “The Effects of Nonnative Accents on Listening Comprehension: Implications for ESL Assessment”. In: *TESOL Quarterly* 36.2, pp. 173–190. ISSN: 1545-7249. DOI: 10.2307/3588329. (Visited on 05/12/2023).
- McDonald, M. et al. (Apr. 2018). “Processing and Comprehension of Accented Speech by Monolingual and Bilingual Children”. In: *Language Learning and Development* 14.2, pp. 113–129. ISSN: 1547-5441. DOI: 10.1080/15475441.2017.1404467. (Visited on 05/12/2023).
- Mesquita, B. et al. (1997). “Culture and Emotion”. In: *Handbook of Cross-Cultural Psychology: Basic Processes and Human Development, Vol. 2, 2nd Ed.* Handbook of Cross-Cultural Psychology. Needham Heights, MA, US: Allyn & Bacon, pp. 255–297. ISBN: 978-0-205-16075-4.
- Monrad-Krohn, G. H. (1947). “The Prosodic Quality of Speech and Its Disorders: (A Brief Survey from a Neurologist’s Point of View)”. In: *Acta Psychiatrica Scandinavica* 22.3-4, pp. 255–269. ISSN: 1600-0447. DOI: 10.1111/j.1600-0447.1947.tb08246.x. (Visited on 05/12/2023).
- Morton, J. B. and Trehub, S. E. (2001). “Children’s Understanding of Emotion in Speech”. In: *Child Development* 72, pp. 834–843. ISSN: 1467-8624. DOI: 10.1111/1467-8624.00318.
- Moskowitz, D. S. (1986). “Comparison of Self-Reports, Reports by Knowledgeable Informants, and Behavioral Observation Data”. In: *Journal of Personality* 54.1, pp. 294–317. ISSN: 1467-6494. DOI: 10.1111/j.1467-6494.1986.tb00396.x. (Visited on 05/12/2023).

- Munro, M. J. and Derwing, T. M. (1995). "Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners". In: *Language Learning* 45.1, pp. 73–97. ISSN: 1467-9922. DOI: 10.1111/j.1467-1770.1995.tb00963.x. (Visited on 05/12/2023).
- Munro, M. J. et al. (Mar. 2006). "The Mutual Intelligibility of L2 Speech". In: *Studies in Second Language Acquisition* 28.1, pp. 111–131. ISSN: 1470-1545, 0272-2631. DOI: 10.1017/S0272263106060049. (Visited on 05/12/2023).
- Murray, I. R. and Arnott, J. L. (1993). "Toward the Simulation of Emotion in Synthetic Speech: A Review of the Literature on Human Vocal Emotion". In: *Journal of the Acoustical Society of America* 93, pp. 1097–1108. ISSN: 0001-4966. DOI: 10.1121/1.405558.
- Niedenthal, P. M. et al. (1994). "Emotional Organization of Perceptual Memory". In: *The Heart's Eye: Emotional Influences in Perception and Attention*. San Diego, CA, US: Academic Press, pp. 87–113. ISBN: 978-0-12-410560-7. DOI: 10.1016/B978-0-12-410560-7.50011-9.
- Öhman, A. and Mineka, S. (2001). "Fears, Phobias, and Preparedness: Toward an Evolved Module of Fear and Fear Learning". In: *Psychological Review* 108, pp. 483–522. ISSN: 1939-1471. DOI: 10.1037/0033-295X.108.3.483.
- Opitz, B. and Degner, J. (July 2012). "Emotionality in a Second Language: It's a Matter of Time". In: *Neuropsychologia* 50.8, pp. 1961–1967. ISSN: 0028-3932. DOI: 10.1016/j.neuropsychologia.2012.04.021. (Visited on 05/12/2023).
- Paap, K. R. et al. (2015). "Bilingual Advantages in Executive Functioning Either Do Not Exist or Are Restricted to Very Specific and Undetermined Circumstances". In: *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior* 69, pp. 265–278. ISSN: 1973-8102. DOI: 10.1016/j.cortex.2015.04.014.
- Pantos, A. J. and Perkins, A. W. (Mar. 2013). "Measuring Implicit and Explicit Attitudes Toward Foreign Accented Speech". In: *Journal of Language and Social Psychology* 32.1, pp. 3–20. ISSN: 0261-927X, 1552-6526. DOI: 10.1177/0261927X12463005. (Visited on 05/12/2023).

- Paulmann, S. and Kotz, S. A. (Jan. 2008). “Early Emotional Prosody Perception Based on Different Speaker Voices”. In: *NeuroReport* 19.2, p. 209. ISSN: 0959-4965. DOI: 10.1097/WNR.0b013e3282f454db. (Visited on 05/12/2023).
- Paulmann, S. and Pell, M. D. (2011). “Is There an Advantage for Recognizing Multi-Modal Emotional Stimuli?” In: *Motivation and Emotion* 35, pp. 192–201. ISSN: 1573-6644. DOI: 10.1007/s11031-011-9206-0.
- Paulmann, S. and Uskul, A. K. (2014). “Cross-Cultural Emotional Prosody Recognition: Evidence from Chinese and British Listeners”. In: *Cognition & Emotion* 28.2, pp. 230–244. ISSN: 1464-0600. DOI: 10.1080/02699931.2013.812033.
- Paulmann, S. et al. (Mar. 2008). “How Aging Affects the Recognition of Emotional Speech”. In: *Brain and Language* 104.3, pp. 262–269. ISSN: 0093-934X. DOI: 10.1016/j.bandl.2007.03.002. (Visited on 05/12/2023).
- Paulmann, S. et al. (Mar. 2011). “Emotional Speech Perception Unfolding in Time: The Role of the Basal Ganglia”. In: *PLOS ONE* 6.3, e17694. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0017694. (Visited on 05/12/2023).
- Paulmann, S. et al. (2013). “Valence, Arousal, and Task Effects in Emotional Prosody Processing”. In: *Frontiers in Psychology* 4. ISSN: 1664-1078. (Visited on 05/12/2023).
- Pavlenko, A. (June 2004). “‘Stop Doing That, Ia Komu Skazala!’: Language Choice and Emotions in Parent—Child Communication”. In: *Journal of Multilingual and Multicultural Development* 25.2-3, pp. 179–203. ISSN: 0143-4632. DOI: 10.1080/01434630408666528. (Visited on 05/12/2023).
- (2005). *Emotions and Multilingualism*. Emotions and Multilingualism. New York, NY, US: Cambridge University Press, pp. xiv, 304. ISBN: 978-0-521-84361-4. DOI: 10.1017/CB09780511584305.
 - (2008). “Emotion and Emotion-Laden Words in the Bilingual Lexicon”. In: *Bilingualism: Language and Cognition* 11, pp. 147–164. ISSN: 1469-1841. DOI: 10.1017/S1366728908003283.

- Pavlenko, A. (2012). "Affective Processing in Bilingual Speakers: Disembodied Cognition?" In: *International Journal of Psychology* 47, pp. 405–428. ISSN: 1464-066X. DOI: 10.1080/00207594.2012.743665.
- Pell, M. D. (2001). "Influence of Emotion and Focus Location on Prosody in Matched Statements and Questions". In: *Journal of the Acoustical Society of America* 109, pp. 1668–1680. ISSN: 0001-4966. DOI: 10.1121/1.1352088.
- (Apr. 2005). "Nonverbal Emotion Priming: Evidence from the 'Facial Affect Decision Task'". In: *Journal of Nonverbal Behavior* 29.1, pp. 45–73. ISSN: 1573-3653. DOI: 10.1007/s10919-004-0889-8. (Visited on 05/12/2023).
- Pell, M. D. and Kotz, S. A. (2011). "On the Time Course of Vocal Emotion Recognition". In: *PLoS ONE* 6. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0027256.
- Pell, M. D. et al. (2009a). "Factors in the Recognition of Vocally Expressed Emotions: A Comparison of Four Languages". In: *Journal of Phonetics* 37, pp. 417–435. ISSN: 1095-8576. DOI: 10.1016/j.wocn.2009.07.005.
- Pell, M. D. et al. (2009b). "Recognizing Emotions in a Foreign Language". In: *Journal of Nonverbal Behavior* 33, pp. 107–120. ISSN: 1573-3653. DOI: 10.1007/s10919-008-0065-7.
- Pinet, M. et al. (Sept. 2011). "Second-Language Experience and Speech-in-Noise Recognition: Effects of Talker–Listener Accent Similarity". In: *The Journal of the Acoustical Society of America* 130.3, pp. 1653–1662. ISSN: 0001-4966. DOI: 10.1121/1.3613698. (Visited on 05/12/2023).
- Ponari, M. et al. (2015). "Processing Advantage for Emotional Words in Bilingual Speakers". In: *Emotion* 15, pp. 644–652. ISSN: 1931-1516. DOI: 10.1037/emo0000061.
- Romero-Rivas, C. et al. (Mar. 2015). "Processing Changes When Listening to Foreign-Accented Speech". In: *Frontiers in Human Neuroscience* 9, p. 167. ISSN: 1662-5161. DOI: 10.3389/fnhum.2015.00167. (Visited on 06/16/2023).

- Rosenthal, R. (1963). "On the Social Psychology of the Psychological Experiment: 1, 2 the Experimenter's Hypothesis as Unintended Determinant of Experimental Results". In: *American Scientist* 51.2, pp. 268–283. ISSN: 0003-0996. JSTOR: 27838693. (Visited on 05/12/2023).
- Rosenthal, R. et al. (1979). "Measuring Sensitivity to Nonverbal Communication: The PONS Test". In: *Nonverbal Behavior*. Elsevier, pp. 67–98.
- Russell, J. A. (Jan. 1994). "Is There Universal Recognition of Emotion from Facial Expression? A Review of the Cross-Cultural Studies". In: *Psychological Bulletin* 115.1, pp. 102–141. ISSN: 0033-2909. DOI: 10.1037/0033-2909.115.1.102.
- Russell, J. A. (1993). "Forced-Choice Response Format in the Study of Facial Expression". In: *Motivation and Emotion* 17, pp. 41–51. ISSN: 1573-6644. DOI: 10.1007/BF00995206.
- Ryan, E. B. (1983). "Social Psychological Mechanisms Underlying Native Speaker Evaluations of Non-Native Speech". In: *Studies in Second Language Acquisition* 5.2, pp. 148–159. ISSN: 1470-1545, 0272-2631. DOI: 10.1017/S0272263100004824. (Visited on 05/12/2023).
- Sauter, D. A. et al. (Feb. 2010a). "Cross-Cultural Recognition of Basic Emotions through Nonverbal Emotional Vocalizations". In: *Proceedings of the National Academy of Sciences* 107.6, pp. 2408–2412. DOI: 10.1073/pnas.0908239106. (Visited on 05/12/2023).
- Sauter, D. A. et al. (2010b). "Perceptual Cues in Nonverbal Vocal Expressions of Emotion". In: *The Quarterly Journal of Experimental Psychology* 63, pp. 2251–2272. ISSN: 1747-0226. DOI: 10.1080/17470211003721642.
- Scherer, K. R. and Wallbott, H. G. (1994). "Evidence for Universality and Cultural Variation of Differential Emotion Response Patterning". In: *Journal of Personality and Social Psychology* 66, pp. 310–328. ISSN: 1939-1315. DOI: 10.1037/0022-3514.66.2.310.
- Scherer, K. R. et al. (1991). "Vocal Cues in Emotion Encoding and Decoding". In: *Motivation and Emotion* 15, pp. 123–148. ISSN: 1573-6644. DOI: 10.1007/BF00995674.

- Scherer, K. R. et al. (2001). “Emotion Inferences from Vocal Expression Correlate across Languages and Cultures”. In: *Journal of Cross-Cultural Psychology* 32, pp. 76–92. ISSN: 1552-5422. DOI: 10.1177/0022022101032001009.
- Schirmer, A. and Kotz, S. A. (Nov. 2003). “ERP Evidence for a Sex-Specific Stroop Effect in Emotional Speech”. In: *Journal of Cognitive Neuroscience* 15.8, pp. 1135–1148. ISSN: 0898-929X. DOI: 10.1162/089892903322598102. (Visited on 05/12/2023).
- (2006). “Beyond the Right Hemisphere: Brain Mechanisms Mediating Vocal Emotional Processing”. In: *Trends in Cognitive Sciences* 10, pp. 24–30. ISSN: 1879-307X. DOI: 10.1016/j.tics.2005.11.009.
- Schirmer, A. et al. (Aug. 2002). “Sex Differentiates the Role of Emotional Prosody during Word Processing”. In: *Cognitive Brain Research* 14.2, pp. 228–233. ISSN: 0926-6410. DOI: 10.1016/S0926-6410(02)00108-8. (Visited on 05/12/2023).
- Schirmer, A. et al. (Apr. 2005). “Sex Differences in the Preattentive Processing of Vocal Emotional Expressions”. In: *NeuroReport* 16.6, p. 635. ISSN: 0959-4965. (Visited on 05/12/2023).
- Schirmer, A. et al. (Mar. 2013). “Vocal Emotions Influence Verbal Memory: Neural Correlates and Interindividual Differences”. In: *Cognitive, Affective, & Behavioral Neuroscience* 13.1, pp. 80–93. ISSN: 1531-135X. DOI: 10.3758/s13415-012-0132-8. (Visited on 05/12/2023).
- Shochi, T. et al. (2010). “Learning Effect of Prosodic Social Affects for Japanese Learners of French Language”. In: (visited on 05/12/2023).
- Siegal, M. et al. (2009). “Bilingualism and Conversational Understanding in Young Children”. In: *Cognition* 110, pp. 115–122. ISSN: 1873-7838. DOI: 10.1016/j.cognition.2008.11.002.
- Souza, A. L. et al. (2013). “Bilingual and Monolingual Children Prefer Native-Accented Speakers”. In: *Frontiers in Psychology* 4. ISSN: 1664-1078. (Visited on 05/12/2023).

- Spreckelmeyer, K. N. et al. (Jan. 2006). “Combined Perception of Emotion in Pictures and Musical Sounds”. In: *Brain Research* 1070.1, pp. 160–170. ISSN: 0006-8993. DOI: 10.1016/j.brainres.2005.11.075. (Visited on 05/12/2023).
- Spreckelmeyer, K. N. et al. (Feb. 2009). “Neural Processing of Vocal Emotion and Identity”. In: *Brain and Cognition* 69.1, pp. 121–126. ISSN: 1090-2147. DOI: 10.1016/j.bandc.2008.06.003.
- Stibbard, R. M. and Lee, J.-I. (July 2006). “Evidence against the Mismatched Interlanguage Speech Intelligibility Benefit Hypothesis”. In: *The Journal of the Acoustical Society of America* 120.1, pp. 433–442. ISSN: 0001-4966. DOI: 10.1121/1.2203595. (Visited on 05/12/2023).
- Strange, W. (1995). “Speech Perception and Linguistic Experience: Issues in Cross-Language Research”. In: *Workshop on Cross-Language Speech Perception*.
- Strauß, A. et al. (Aug. 2013). “Narrowed Expectancies under Degraded Speech: Revisiting the N400”. In: *Journal of Cognitive Neuroscience* 25.8, pp. 1383–1395. ISSN: 1530-8898. DOI: 10.1162/jocn_a_00389.
- Sun, Y. et al. (Feb. 2022). “How Accentuation Influences the Processing of Emotional Words in Spoken Language: An ERP Study”. In: *Neuropsychologia* 166, p. 108144. ISSN: 0028-3932. DOI: 10.1016/j.neuropsychologia.2022.108144. (Visited on 05/12/2023).
- Sutton, T. M. et al. (Aug. 2007). “The Automatic Access of Emotion: Emotional Stroop Effects in Spanish–English Bilingual Speakers”. In: *Cognition and Emotion* 21.5, pp. 1077–1090. ISSN: 0269-9931. DOI: 10.1080/02699930601054133. (Visited on 05/12/2023).
- Thompson, W. F. and Balkwill, L. L. (2006). “Decoding Speech Prosody in Five Languages”. In: *Semiotica* 158, pp. 407–424. ISSN: 0037-1998. DOI: 10.1515/SEM.2006.017. (Visited on 05/12/2023).
- Tomkins, S. S. and McCarter, R. (Feb. 1964). “What and Where Are the Primary Affects? Some Evidence for a Theory”. In: *Perceptual and Motor Skills* 18.1,

- pp. 119–158. ISSN: 0031-5125. DOI: 10.2466/pms.1964.18.1.119. (Visited on 05/12/2023).
- Van Bezooijen, R. et al. (1983). “Recognition of Vocal Expressions of Emotion: A Three-Nation Study to Identify Universal Characteristics”. In: *Journal of Cross-Cultural Psychology* 14, pp. 387–406. ISSN: 1552-5422. DOI: 10.1177/0022002183014004001.
- van den Noort, M. et al. (Mar. 2019). “Does the Bilingual Advantage in Cognitive Control Exist and If So, What Are Its Modulating Factors? A Systematic Review”. In: *Behavioral Sciences (Basel, Switzerland)* 9.3, p. 27. ISSN: 2076-328X. DOI: 10.3390/bs9030027.
- Wallbott, H. G. and Scherer, K. R. (1986). “Cues and Channels in Emotion Recognition”. In: *Journal of Personality and Social Psychology* 51, pp. 690–699. ISSN: 1939-1315. DOI: 10.1037/0022-3514.51.4.690.
- Wambacq, I. J. A. and Jerger, J. F. (Aug. 2004). “Processing of Affective Prosody and Lexical-Semantics in Spoken Utterances as Differentiated by Event-Related Potentials”. In: *Cognitive Brain Research* 20.3, pp. 427–437. ISSN: 0926-6410. DOI: 10.1016/j.cogbrainres.2004.03.015. (Visited on 05/12/2023).
- Wei, H. et al. (May 2022). “An EEG-study on L2 Categorization of Emotional Prosody in German”. In: *Speech Prosody 2022*, pp. 629–633. DOI: 10.21437/SpeechProsody.2022-128. (Visited on 05/12/2023).
- Weimer, A. et al. (2022). “Behavioral Evidence for Differences in Emotional Resonance during Processing First and Second Language”. In: *International Journal of Bilingualism* 26.6, pp. 784–798.
- Wells, J. C. and Wells, J. C. (1982). *Accents of English: Volume 1*. Vol. 1. Cambridge University Press.
- Wilson, D. and Wharton, T. (Oct. 2006). “Relevance and Prosody”. In: *Journal of Pragmatics*. Special Issue: Prosody and Pragmatics 38.10, pp. 1559–1579. ISSN: 0378-2166. DOI: 10.1016/j.pragma.2005.04.012. (Visited on 05/12/2023).

- Witteman, J. et al. (Nov. 2011). “The Nature of Hemispheric Specialization for Linguistic and Emotional Prosodic Perception: A Meta-Analysis of the Lesion Literature”. In: *Neuropsychologia* 49.13, pp. 3722–3738. ISSN: 1873-3514. DOI: 10.1016/j.neuropsychologia.2011.09.028.
- Yan, Y. et al. (Dec. 2022). “In-Group Advantage for Chinese and English Emotional Prosody in Quiet and Noise Conditions”. In: *2022 13th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, pp. 305–309. DOI: 10.1109/ISCSLP57327.2022.10037894.
- Yow, W. Q. and Markman, E. (2007). “Monolingual and Bilingual Children’s Use of Mutual Exclusivity Assumption and Pragmatic Cues in Word Learning”. In: *Poster Presented at the Meeting of the Society for Research in Child Development, Boston, MA*.
- (2009). “Understanding a Speaker’s Communicative Intent: Bilingual Children’s Heightened Sensitivity to Referential Gestures”. In: *Proceedings of the 33rd Annual Boston University Conference on Language Development*. Vol. 2, pp. 646–657.
 - (2011). “Bilingualism and Children’s Use of Paralinguistic Cues to Interpret Emotion in Speech”. In: *Bilingualism: Language and Cognition* 14, pp. 562–569. ISSN: 1469-1841. DOI: 10.1017/S1366728910000404.
- Zhu, Y. (Jan. 2013). “Which Is the Best Listener Group?: Perception of Chinese Emotional Prosody by Chinese Natives, Naïve Dutch Listeners and Dutch L2 Learners of Chinese”. In: *Dutch Journal of Applied Linguistics* 2.2, pp. 170–183. ISSN: 2211-7245, 2211-7253. DOI: 10.1075/dujal.2.2.03zhu. (Visited on 05/12/2023).
- Zuckerman, M. et al. (1975). “Encoding and Decoding Nonverbal Cues of Emotion”. In: *Journal of Personality and Social Psychology* 32, pp. 1068–1076. ISSN: 1939-1315. DOI: 10.1037/0022-3514.32.6.1068.

Appendix A

Stimuli

Final Stimuli (Turkish)

- Mutfak havlusu katlandı.
- Elbise askıya asıldı.
- O güvertede durur ve gazete okur.
- O köşede bekler ve kitap okur.
- O marketin karşısında bekler.
- Onun kitabı yatağının altında.
- O bardağı masanın üstüne koydu.
- Benim kaşığım masanın üstünde.
- Bu çatal gümüşten yapılmıştır.
- Onun kamerası çantasında.
- Onun defteri çatıda.
- O akademik bir makale yazdı.

- Dünya kabaca küresel.
- Erkek geyik türlerinin boynuzları çıkar.
- Güneş sisteminde 8 gezegen vardır.
- Dünya Güneşin uydusudur.
- Magnetler buzdolabının üstünde.
- Bu masa ahşaptan yapıldı.
- Bu masa mozaikle kaplı.
- Bu koltuk yeşil kadife ile kaplı.
- Bu perde ipekten yapıldı.
- Boynuzlar hayvanların kafatasının uzantılarıdır.
- Birçok dişi ren geyiğinin boynuzu vardır.
- Su 100 derecede kaynar.
- Çello 4 telli bir çalgıdır.
- Nijerya Batı Afrika'da bulunur.
- Dijital saatler yaygındır.
- Yerde bir İran halısı gördüm.
- Bu İran halısı yerli insanlar tarafından yapıldı.
- Bizim vücudumuz sudan oluşur.
- Su 0 derecenin altında buz olur.
- Bu bina 20 katlıdır.
- Odada çok fazla kutu var.
- Bu bir metal çöp tenekesi.

- Dolapta 4 çekmece var.
- Konteynerların mavi bir kapağı var.
- Bu şişe camdan yapılmış.
- Turuncu pipetler metalikler.
- Bu bıçaklar metalden yapılmış.
- Tabak yuvarlak masanın üzerinde.
- Müşteriler bir çift şort aldı.
- Bu masa metalle kaplı.
- O mavi kanvas bir gömlek satın aldı.
- Bu kapı ahşaptan yapılmış.
- Pencereleler gazete ile temizlenir.
- Bu glutensiz kayısı reçeli.
- Türkiye, Asya ile Avrupa arasında bir köprüdür.
- Maymunlarla benzer genleri paylaşıyoruz
- Yetişkin filler zıplayamaz
- Müşteri bir kitap ve dergi satın aldı
- O yıllarca muhasebecilik yaptı
- O kafede garson
- O okulun restoranında aşçıdır.
- O şekersiz kek yaptı
- Tokyo dünyanın en kalabalık şehridir
- İtalya'nın 450 adası var.

- Türkiye’de 2 ada var
- O deri cüzdan aldı.
- Evde 7 tablo var
- Duvarlar aynalarla kaplıdır.
- Kaşık ve çatal masanın üzerinde
- Kurbağalar yeşildir ve noktaları vardır
- Bu 3 kapılı bir dolaptır
- Finlandiya’da 6 ay geceder.
- Finlandiya’da 6 ay gündüzdür.
- Sığırlar meyve ve sebze yerler.
- Birçok karga türü siyah veya gridir.
- Bahçede 13 farklı çiçek var
- Türkiye’de 4 mevsim vardır
- Türkiye üç tarafı denizlerle çevrilidir.
- Türkiye’nin üç tarafı denizlerle kaplıdır.
- Bu mavi kuşun beyaz noktaları var.
- Bu kazak yünden yapılmıştır
- Bu film bir kitap uyarlamasıdır.
- O kitabı kitaplığa koydu.
- Penguenler suda yaşayan uçamayan kuşlardır.
- Hanımeli sarı bir çiçektir
- Bu el yapımı bir mavi halı

- O şapka ve gözlük takıyor
- Onun yeşil gözleri ve kahverengi saçları var
- Bu kırmızı çaydanlık seramikten yapılmıştır.
- Kahve çekirdekleri Kolombiya'dan geliyor
- Türk çayı yaprakları Rize'de yetiştiriliyor.
- Karamel eritilmiş şekerden yapılır
- O havluları dolaba koydu
- O tişörtleri ve pantolonları katlıyor.
- İki pantolon aldım.
- Yiyecekler mutfak dolaplarında saklanır
- Anahtarlar mutfak dolabında
- Dolaptan ütü masasını çıkardı.
- Dolapta bir kutu gördü
- Buzdolabından bir şişe su aldı
- İlk dönme dolap 1893'te yapıldı
- Kitaplar kağıttan yapılır
- Televizyonda maç izleyen iki genç adam gördüm
- Televizyonda bir belgesel izliyor.
- Kanepeye turuncu minderler koydu.

Final Stimuli (English)

- A kitchen towel is folded.
- The dress is hung on a hanger.
- He stands on the deck and reads a newspaper.
- She waits on the corner and reads a book.
- He waits across the market.
- Her book is under her bed.
- He put the glass on the table.
- My spoon is on the table.
- This fork is made of silver.
- Her camera is in the bag.
- His notebook is on the roof.
- She wrote an academic article.
- The earth is roughly spherical.
- Male deer species grow antlers.
- There are 8 planets in the solar system.
- Earth is a satellite of the sun.
- There are magnets on the fridge.
- This table is made of wood.
- This table is covered with mosaic.
- This sofa is covered with green velvet.

- This curtain is made of silk.
- Antlers are extensions of an animal's skull.
- Most female caribou have antlers.
- Your music sheets are on the stand.
- The cello is a 4-string instrument.
- Nigeria is located in West Africa
- I see a Persian rug on the floor.
- This Persian rug is handmade.
- Our body is made of water.
- More than half of the Earth is made up of water.
- The largest ocean is the Pacific Ocean
- The highest mountain is Mount Everest
- There are lots of bins in the room
- This is a metal garbage can.
- There are four drawers in the cabinet.
- Containers have a blue lid.
- This bottle is made of glass.
- Orange pipes are metallic
- These knives are made of metal
- The plate is on the round table
- The customers bought a pair of shorts
- This table is covered with metal

- He bought a blue canvas shirt
- This door is made of wood
- Windows are cleaned with a newspaper
- This is gluten free apricot jam
- Turkey forms a land bridge between Asia and Europe.
- We share similar genes with monkeys
- Adult elephants cannot jump
- The customer bought a book and magazines
- She was an accountant for years
- He is a waiter at the cafe
- She is a cook at the school restaurant
- He made a sugar free cake
- Tokyo is the most crowded city in the world
- Italy has 450 islands.
- There are 2 islands in Turkey
- He bought a leather wallet.
- There are 7 paintings in the house
- Walls are covered with mirrors.
- Spoon and fork are on the table
- Frogs are green and have dots
- This is a 3-door cabinet
- It is a 6 months night in Finland

- It is daytime for 6 months in Finland.
- Cattle eat fruits and vegetables.
- Many species of crows are black or grey.
- There are 13 different flowers in the garden
- There are 4 seasons in Turkey
- Turkey is surrounded on three sides by the sea.
- Three sides of Turkey are covered by seas.
- This blue bird has white dots.
- This jumper is made of wool
- This movie is a book adaptation.
- She put the book on the bookshelf.
- Penguins are aquatic flightless birds
- Honeysuckle is a yellow flower
- This is a handmade blue carpet
- She wears a hat and glasses
- He has green eyes and brown hair
- This red teapot is made of ceramic
- Coffee beans come from Colombia
- Caramel is made of melted sugar
- He put towels in the cupboard
- I bought two pairs of trousers.
- Tom made a pot of coffee

- Food and cookware are stored in kitchen cabinets
- The keys are in the kitchen wardrobe
- She pulled the ironing board out of the closet
- He saw a box in the closet
- She got a bottle of water out of the fridge
- The first Ferris wheel was built in 1893
- Books are made out of paper
- I saw two young men watching a match on television
- He is watching a documentary on television
- He put orange cushions on the sofa.

Appendix B

Language History Questionnaire



Language History Questionnaire

1. Participant no

2. Gender

Mark only one oval.

- ☐ Male
- ☐ Female
- ☐ Non-binary/ third gender
- ☐ Prefer not to say

3. Age (in years)

4. Handedness

Mark only one oval.

- ☐ left
- ☐ right

5. Native country:

6. Do you have any known visual or hearing problems (corrected or uncorrected)?

Mark only one oval.

☐ Yes

☐ No

7. If you selected 'yes', please elaborate (e.g., you wear glasses/contacts, hearing aids, etc.).

8. 1) What is your first language (i.e., language first spoken)? If more than one, please briefly describe the situations in which each language was used.

9. 2) Which language (if any) do you consider your second language?

10. 3) If you have ever lived in or visited a country where languages other than your native language are spoken, please indicate below the name of the country (countries), the duration of your stay in number of months (if less than 4 weeks, do not include), and which languages you used while you were in the country (please indicate if you were spoken to in a language other than your first language, even if you never actually spoke that language).

11. 4) List below, from most fluent to least fluent, all of the languages to which you have been exposed. Also specify the age in years at which you began to learn the language and the context in which you learned it. For example, "Turkish, birth, home". Include all languages to which you have been exposed, although you may never have had formal training in them and may not be able to read, speak or write them.

12. 5) What languages were spoken in your home while you were a child and by whom?

13. 6-a) Please list the language(s) the following people speak. Parent(s)

14. 6-b) Please list the language(s) the following people speak, Other caregiver(s),

15. 6-c) Please list the language(s) the following people speak, Closest friend:

16. 7) How many years have you studied your second language? Please indicate the setting(s) in which you have had experience with the language (i.e., classroom, with friends, foreign country...)

17. 8) Are you currently enrolled in any language courses? (Any course either instructed in a foreign language or designed to teach a foreign language). If so, please list the course number(s) below, along with the title of the course.

18. 9) What languages other than your first language do you speak fluently?

- Mark only one oval.

not | ○ ○ ○ ○ ○ ○ ○ ○ very literate

23. 14) Please rate your second language reading proficiency on a ten-point scale.(1 = not literate, 10 = very literate)

Mark only one oval.

[illegible]

24. 15) Please rate your first language writing proficiency on a ten-point scale.(1 = not literate, 10 = very literate)

Mark only one oval.

1 2 3 4 5 6 7 8 9 10

not | ○ ○ ○ ○ ○ ○ ○ ○ ○ very literate

25. 16) Please rate your second language writing proficiency on a ten-point scale.(1 = not literate, 10 = very literate)

Mark only one oval.

[illegible]

26. 17) Please rate your first language conversational fluency on a ten-point scale.(1 = not fluent, 10 = very fluent)

Mark only one oval.

[illegible]

27. 18) Please rate your second language conversational fluency on a ten-point scale.(1 = not fluent, 10 = very fluent)

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
not f	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	very fluent

28. 19) Please rate your first language speech comprehension ability on a ten-point scale.(1 = unable to understand conversation, 10 = perfectly able to understand conversation)

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
no c	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	perfect comprehension

29. 20) Please rate your second language speech comprehension ability on a ten-point scale.(1 = unable to understand conversation, 10 = perfectly able to understand conversation)

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
no c	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	perfect comprehension

30. How would you rate your foreign language learning skills?

Mark only one oval.

- ☐ worse than average
- ☐ average
- ☐ better than average

31. 22) When learning a foreign language, which of the following do you find the easiest to learn? Please rank the following from 1 to 4. (1 = easiest, 4 =hardest)

Mark only one oval per row.

	Pronunciation	Grammar	Vocabulary	Sayings/Expressions
1	☐	☐	☐	☐
2	☐	☐	☐	☐
3	☐	☐	☐	☐
4	☐	☐	☐	☐

32. 23) Have you ever been immersed in your second language culture (e.g., living abroad)?

Mark only one oval.

☐ Yes

☐ No

33. 24) Please compare learning a second language in an immersion environment with learning one in a classroom environment. (E.g., Which is easier? In which did you learn more?...). Please comment about the difference between your various learning experiences.

34. 24) Please compare learning a second language in an immersion environment with learning one in a classroom environment. (E.g., Which is easier? In which did you learn more?...). Please comment about the difference between your various learning experiences.

This content is neither created nor endorsed by Google.

Google Forms

Appendix C

EEG Pre-Screening Form



EEG Pre-Screening Form

1. Participant No

2. Date of Birth

Example: 7 January 2019

3. Gender:

Mark only one oval.

- ☐ Male
- ☐ Female
- ☐ Non-binary/ third gender
- ☐ Prefer not to say

4. Handedness:

Mark only one oval.

- ☐ Left
- ☐ Right

5. Normal hearing?

Mark only one oval.

- ☐ Yes
- ☐ No

6. Normal (or corrected) vision?

Mark only one oval.

- ☐ Normal Vision
- ☐ Corrected Vision (Glasses, Contacts etc.)

7. Native Language

8. Do you have any psychiatric or neurological disorder (e.g., epilepsy)?

Mark only one oval.

- ☐ Yes
- ☐ No

9. Do you use psychiatric medication?

Mark only one oval.

- ☐ Yes
- ☐ No

10. Do you have any metallic implants around your neck or head?

Mark only one oval.

- ☐ Yes
- ☐ No

11. Are you pregnant or could you possibly be pregnant?

Mark only one oval.

☐ Yes

☐ No

12. How many hours do you usually sleep?

13. How many hours did you sleep last night?

This content is neither created nor endorsed by Google.

Google Forms