IN SEARCH OF LANGUAGE-BASED FACTORS INFLUENCING RHYTHMIC GROUPING

ECE KAYA

BOĞAZİÇİ UNIVERSITY

2019

IN SEARCH OF LANGUAGE-BASED FACTORS

INFLUENCING RHYTHMIC GROUPING

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Cognitive Science

by

Ece Kaya

Boğaziçi University

2019

DECLARATION OF ORIGINALITY

I, Ece Kaya, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature..

Date ....15.05.2019..

ABSTRACT

In Search of Language-Based Factors Influencing Rhythmic Grouping

The Iambic-Trochaic Law (Hayes, 1995) was proposed as a universal mechanism governing rhythmic grouping. It predicts that sequences of sound events are grouped trochaically (strong-weak) when they alternate in intensity or pitch, and iambically (short-long) when they alternate in duration (Langus et al., 2016). Previous studies have found native language effects on grouping of both linguistic (Crowhurst & Teodocio Olivares, 2014; Langus et al., 2016) and non-linguistic (Iversen et al., 2008; Molnar et al., 2016) sounds, as they revealed grouping preferences incompatible with the predictions of the ITL. The present study investigated native language effects on rhythmic grouping by presenting native Turkish speakers sequences of tones (Experiment 1A and Experiment 1B) and syllables (Experiment 2). Lack of native language effects on grouping as revealed by the results, was explained by the stress 'deafness' (i. e., reduced sensitivity to stress deviations) exhibited by Turkish speakers, under the assumption that Turkish has predictable, fixed stress. That speakers of languages with fixed, lexically unmarked stress were reported to exhibit similar insensitivities to stress deviations (Dupoux et al., 1997; Peperkamp & Dupoux, 2002; Rahmani et al., 2015; Lu et al., 2018) supported this conclusion.

ÖZET

Ritmik Gruplamada Anadil Temelli Etkilerin Arayışı

Iambik-Trokaik Yasa (Hayes, 1995), ritmik gruplamanın altında yatan evrensel bir mekanizma olarak öne sürülmüştür. Bu yasaya göre, ses düzeyi veya ses perdesi bakımından değişen ikili ses grubundan oluşan bir dizideki sesler, 'trokaik' olarak, yani kuvvetli olan önce gelecek şekilde gruplanır (Langus et al., 2016). Süre bakımından değişen ikili ses grubundan oluşan bir dizideki sesler ise 'iambik' olarak, yani kısa olan ses önce gelecek şekilde gruplanır (Hayes, 1995). Bu konuda yapılmış olan çalışmalarda, konuşma seslerinden (Crowhurst & Teodocio Olivares, 2014; Langus ve ark., 2016) ve sentetik seslerden (Iversen ve ark., 2008; Molnar ve ark., 2016) oluşan dizilerin gruplanmasında iambik-trokaik yasaya uymayan sonuçlar elde edilmiş, ve bu sonuçlar, anadilin gruplamaya etkisi olarak açıklanmıştır. Mevcut çalışmada, anadili Türkçe olan katılımcılara sentetik seslerden (Deney 1A ve Deney 1B) ve hecelerden (Deney 2) oluşan diziler dinletip gruplama tercihleri incelenmiştir. Gruplamada anadilin etkisi olmadığını gösteren sonuçlar, anadili Türkçe olan kişilerin, Türkçe'deki vurgunun tahmin edilebilir olması nedeniyle, vurgu değişimlerine karşı algısal bir hassasiyet geliştirmemiş olmaları ile açıklanmıştır. Anadilindeki vurgu tahmin edilebilir olan kişilerin, vurgu değişimlerine karşı düşük hassasiyetleri olduğunu raporlayan çalışmalar (Dupoux ve ark., 1997; Peperkamp & Dupoux, 2002; Rahmani ve ark., 2015; Lu ve ark., 2018) bu sonucu destekler niteliktedir.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDIX TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Firstly, I want to thank my advisors; Esra Mungan, for her encouragement, guidance and support not only during the thesis process but also through my Master's journey. Pavel Logacev, for helping me understand statistics from a different perspective.

I am also thankful to Sumru Özsoy and to thesis committee members, İnci Ayhan, Gaye Soley and Ozan Baysal for their valuable comments and advice from different disciplines and points of view which broadened my scope and helped me advance this thesis. I also thank Aniruddh Patel and John Iversen for sharing the materials they used in their study.

I would like to thank Elif Canseza Kaplan and Duygu Özbağcı for their friendship and support, which no distance could hinder; and Emre Oral for his help on coding and his company at the lab.

I also thank my family, who always supports me in all aspects, and to my grandfather, Erol Kaya for commencing my interest in linguistics at a young age.

This thesis is a part of the project "In Search of Influences of Music and Language on Rhythm Perception" supported by the Boğaziçi University Scientific Research Fund Grant No. 14324. I thank Boğaziçi University for funding the project and providing me with research opportunities.

To Leon and Mathilda,

CHAPTER 1

INTRODUCTION

1.1 Rhythm perception and rhythmic grouping

Auditory processing involves perceiving and segmenting the sound signal and carrying it to higher levels of cognition. This process is facilitated by grouping low-level elements into units, which gives rise to the sensation of rhythm or melodic motive, for instance. Regardless of the structure, rhythm can be conceived as a product of grouping (Hunyadi, 2015). Humans tend to organize the auditory stream into rhythmic units even when no physical cues are present. This phenomenon was first observed by Bolton, who found that listeners perceive isochronous and identical sound pulses (or: sound clicks) as if they were of two kinds alternating with each other as “tick-tock" sequence (Bolton, 1894). This tendency for grouping constitutes the core of rhythm cognition.

The definition of rhythm varies across disciplines and the inquiry in which it is centered. It is usually conceptualized as patterns of sound in speech or music but it can also be interpreted beyond sound, for instance in the domain of vision. It can also be associated with non-physical objects such as occurrences of events and states of mind (Lefebvre, 2004). In a broad sense, any stream of events which incorporate alternation and repetition can be considered as rhythm. Hunyadi refers to rhythm as “the systematic occurrence of elements and features at distinct spaces, intervals or qualities (Hunyadi, 2015). Fraisse defines rhythm as the repetition of patterned sequences of elements that often vary in prominence (Fraisse, 1974,1982, as cited in Hay & Diehl, 2007). Within the scope of the current study, this latter, narrower conceptualization is used.

The focus of the current study is rhythmic grouping biases in the auditory domain, specifically language. Speech rhythm has an important role in language acquisition (Jusczyk & Aslin, 1995). Developmental research suggests that infants use acoustic cues as a bootstrap mechanism in speech segmentation (Saffran, Aslin & Newport 1996). Langus et al. (2017) states that infants use language rhythm also for the acquisition of the words and word order of their language. The segmentation of speech is facilitated by the stress patterns of the language both at acquisition and in word recognition (Höhle et al. 2009). Speech rhythm forms a core communicative component of spoken language by its beneficial roles in semantics, comprehension and communication (Kotz et al., 2018).

There are various definitions of language rhythm, each of which lead to different considerations and interpretations. Mirroring the hierarchical structure of language, linguistic rhythm is explained and investigated in multiple levels. (Langus et al., 2017). At the lowest level, which is conceptualized as segmental, rhythm is signaled through the alternation between vowels and consonants. Rhythm can be measured through the vocalic space percentage in the speech stream and the standard deviation of consonantal intervals (Ramus et al., 1999) at this level. Above this level, linguistic rhythm is signaled through the alternation of stressed and unstressed elements. According to Langus et al. (2017), speech rhythm at the lower levels contributes to word segmentation and acquisition of words whereas the higher level is crucial in the acquisition of word order.

The differences between languages in terms of rhythm gave rise to the distinction of rhythm classes. Early studies focusing on language rhythm proposed that the world's languages are categorized into rhythm classes defined as stress-timed, syllable-timed and mora-timed (Pike, 1945; Abercrombie, 1967, as cited in Langus et al. 2017). From this classic theory of speech rhythm, it follows that speech signal is divided into equal intervals at any of the levels of linguistic rhythm. Ramus et al. (1999) conveys a body of research that reveal evidence against isochrony and argue that the subjective perception of isochrony must be based on a more abstract construct. In their study, they measure the properties of vowels such as the percentage of vocalic space (%V) and the standard deviation of the consonantal intervals (ΔC) and compare these values across eight languages. Nespor et al. (2011) added six languages to their analysis and reported the differences in their ΔC and %V values, as represented in Figure 1. The rhythmic class of a language appear to be a crucial factor determining the segmentation unit used by its speakers, as supported by the observation that infants segment languages on the basis of their rhythmic class and that this ability may be utilized during language acquisition (Nespor et al., 2011).

Despite decade-long debates, the mechanisms that govern rhythm perception are not well understood. A body of research have revealed that domain-general mechanisms underlie rhythmic grouping in the domains of language (Hay & Diehl, 2007), music (Bolton, 1894) and vision (Langus et al., 2016). A line of research suggests the opposite, that rhythmic grouping biases are influenced by experience, specifically, by culture and native language (Iversen et al., 2008; Schmidt-Kassow et al., 2011; Bhatara et al., 2013; Molnar et al., 2016; Langus et al., 2016; Crowhurst, 2018).

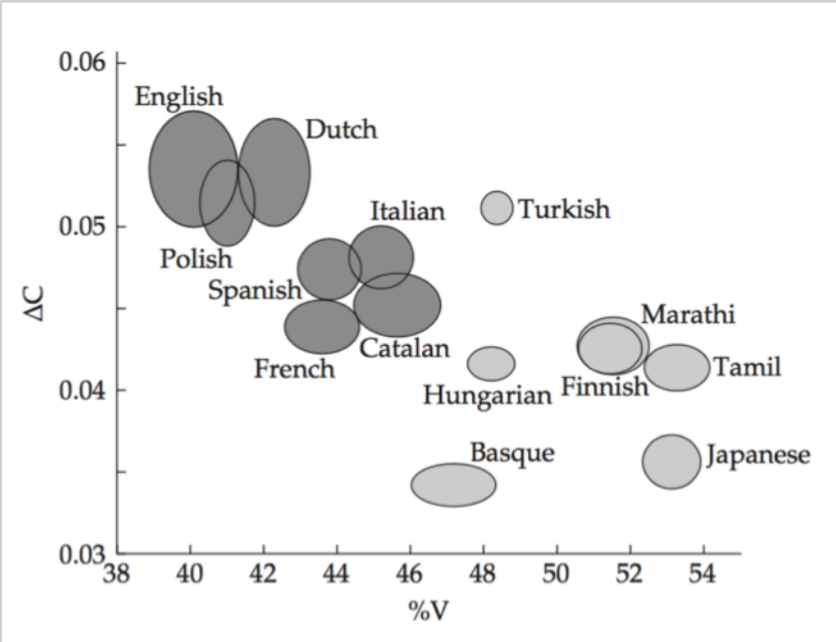

Figure 1 Standard deviation of the consonantal intervals(ΔC) and the percentage of vocalic space (%V) for 14 languages. The widths of the ellipses along the two axes represent standard errors of the mean along the axes. Dark ellipses represent head-initial languages, and light ellipses head-final languages. Figure is from Nespor et al., 2011.

1.1.1 The Iambic – Trochaic Law

The strongest candidate of a domain-general mechanism underlying rhythm perception is the Iambic-Trochaic law (ITL). It was proposed by Hayes (1995), who expanded findings from the auditory grouping studies of Bolton (1894) and Woodrow (1951, as cited in Langus et al., 2016) to the domain of music. The ITL (Hayes, 1995) states that sequences of sounds which alternate in loudness are grouped such that the louder sound precedes the softer sound, forming a “trochee”. Sequences of sounds which alternate in duration are grouped such that the longer sound follows the shorter sound, forming an “iamb”. The terms trochee and iamb refer to strong-weak (Xx) and weak-strong (xX) grouping, respectively, regardless of the parameter which is alternating (e. g., loudness, duration, or pitch). Nespor et al.'s (2008) formulation of the ITL includes pitch, proposing that sequences alternating in pitch are also grouped trochaically. This leads to the following formulation of the ITL (Langus et al., 2016):

1. Elements alternating in intensity or pitch are grouped trochaically, with the strong element (with higher intensity or pitch) preceding the weak one.
2. Elements alternating in duration are grouped iambically, with the weak element (with shorter duration) preceding the strong one (p. 1128).

ITL may be the mechanism which helps infants segment the words and the word order of their language (Bion et al., 2011, Hay & Saffran, 2012, Yoshida et al., 2010). However, there is a debate about whether ITL is a universal perceptual mechanism or is influenced by language experience, or is a combination of both. The studies which investigated ITL cross-linguistically reveal conflicting results. Moreover, the extent to which ITL governs rhythmic grouping is not well understood. Whereas a body of research investigating ITL with linguistic and non-linguistic sounds suggest that ITL is a domain-general mechanism whose effects extend to grouping of non-linguistic sounds, another line of research suggest that cross-linguistic differences only emerge with linguistic stimuli.

1.1.2 Studies investigating ITL

Hay & Diehl (2007) compared the grouping preferences of English and French speakers. They used both linguistic and non-linguistic stimuli. In the experiments, the participants were presented with sequences of square-wave tones alternating in either intensity or duration and were asked whether they heard a strong-weak or weak-strong group. The linguistic stimuli consisted of the sequences with the repetitions of the [ga] syllable presented in alternating intensity or duration. No grouping differences were found between the speakers of both languages. The participants grouped both linguistic and non-linguistic stimuli trochaically when they alternated in intensity and iambically when they alternated in duration. The authors interpreted the lack of grouping difference between the speakers of two languages as evidence for the universality of the ITL, given the prosodic dissimilarity between English and French (Molnar et al., 2016).

With a different experimental design, Langus et al. (2016) obtained similar results only for non-linguistic stimuli. They compared the perceptual grouping biases of speakers of Italian, Turkish and Persian for both linguistic and non-linguistic sounds. Presented with sequences consisting of either the syllables [pa], [su], [tu], [ke], [ma], [vi], [bu], [go], [ne], [du] or their sine-wave analogues, the participants were asked whether a specific non-word was in the sequence. The sine-wave analogues of the syllables were synthesized by replacing three formants in speech with sinusoids. The syllables and their sine-wave analogues were manipulated in terms of their duration and pitch such that either of the participants' response revealed a specific perceptual group. The results of the experiment with linguistic stimuli showed cross-linguistic differences for the duration manipulation. The grouping biases of the speakers of three languages reflected the phrasal prominence patterns of the languages: whereas the speakers of Italian grouped duration sequences iambically, speakers of Turkish and Persian grouped them trochaically. However, no such difference occurred in the experiment with sine-wave stimuli. Given the minor difference between the stimuli, these results point to a role of acoustic cues specific to language that lead to language effects. The results can also be interpreted as the extent to which ITL is affected by language experience.

In contrast, there are studies that reveal cross-linguistic differences with non-linguistic stimuli. For instance, Iversen et al. (2008) compared grouping preferences of speakers of English and Japanese. They presented the participant groups the same sequences of square-wave tones, alternating in either duration or amplitude at various levels of manipulation. Whereas both groups grouped the intensity sequences trochaically, grouping preferences differed in duration sequences. Speakers of English grouped these sequences consistent with the predictions of the ITL, namely, iambically. However, speakers of Japanese tended to group the duration sequences more often trochaically than iambically, in contrast to the predictions of the ITL. The authors proposed the differences in word order between the languages as an explanation of the results, pointing out that the function words precede the content words in English and in Japanese the opposite pattern emerges. Even though such an explanation appears far-fetched at first glance, given that syntax is a higher level component of language cognition, it is compatible with the stress patterns of languages. Nespor et al. (2008) argues that the PP-level rhythm is mirrored in the word order of a language and vice versa. Results of Iversen et al. (2008) indicate that the native language effects on rhythmic grouping carry over to other domains, effecting the perception and grouping of lower-level sound patterns.

Another study using non-linguistic stimuli reveals differences between monolinguals and bilinguals with dominance in languages with different word order, indicating the role of language experience at a more precise level (Molnar et al., 2016). In their experiments, Molnar et al. (2016) adapted the method and materials of Iversen et al. (2008), and presented participants with sequences of square-wave tones. They explored the differences in grouping between four participant groups: Spanish monolinguals, Spanish-dominant bilinguals, Basque-proficient bilinguals and Basque-dominant bilinguals. The sequences alternating in intensity were grouped trochaically by all groups. For the sequences alternating in duration, significant differences in grouping appeared only between the Spanish monolinguals and Basque-dominant bilinguals. Spanish monolinguals grouped the sequences as short-long, namely, iambically whereas the grouping preferences of Basque-dominant bilinguals were the opposite. This difference reflects the differences in word order between Spanish and Basque with the former being a SVO language and the latter being a SOV. In another study, sequences of syllables alternating in duration were grouped trochaically by speakers of Native Mexican Spanish, indicating the effect of dialect on rhythmic grouping biases (Crowhurst, 2018).

Further structural differences between the languages seem to affect grouping preferences even when the languages share the same word order. For instance, Bhatara et al. (2013) found differences in the strength of grouping between speakers of French and German. Participants were presented with sequences of syllables, consisting of combinations of vowels /e/, /i/, /o/, /u/ and consonants /b/, /z/, /m/, /l/ with no pauses in between, and were asked whether they heard a strong-weak or weak-strong grouping. Speakers of both German and French revealed groupings compatible with the ITL: sequences of syllables alternating in duration were grouped iambically and the sequences alternating in pitch and intensity were grouped trochaically. However, the grouping preferences of German speakers were stronger than that of the French speakers who grouped somewhat inconsistently. Though these languages have the same word order, their stress patterns differ. Whereas German has a trochaic word-level stress which is signaled by duration, French has no word-level stress. PP-level stress of German is not fixed and is aligned with the lexical stress whereas French has a fixed PP-final stress which is manifested through increased pitch and duration. Moreover, speakers of French have "stress-deafness" (Dupoux et al., 1997). These differences between French and German is reflected in the results of this study, indicating that stress patterns may be a stronger influence on rhythmic grouping than word order. Other studies, too, found that German speakers are more biased towards ITL-compatible groupings whereas the French speakers lack such a strong sensitivity for such grouping biases. Attributing this difference in consistency between German and French speakers to the "stress-deafness" of French speakers, Yeung et al. (2018) investigated whether learning about the prosodic patterns such as word stress modulates rhythmic grouping. They trained French adults on a German-like stress contrast and found that participants who showed better phonological learning made more consistent ITL-like groupings, particularly over duration cues. The authors concluded that phonological learning modulates low-level auditory grouping, indicating the role of prosodic features in rhythmic grouping.

Schmidt-Kassow et al. (2011) have investigated the differences between French and German speakers using neuroimaging to see whether French late learners of German perceive deviations of trochees in German sentences. They presented participants German sentences with either metrical violations (i. e., the violations of the trochaic stress patterns), syntactic violations or a combination of both violations and measured their ERP responses to the sentences. While ERP responses both to syntactic and metrical violations were obtained in German speakers, French bilinguals showed a response only to syntactic violations, which was interpreted as an indicator of their competence in German. Lack of response to metrical violations was interpreted as an outcome of French bilinguals' stress deafness. In the same study, Schmidt-Kassow et al. (2011) compared ERP responses of both groups to deviations in non-linguistic stimuli, adapting the paradigm used in previous studies on the "tick-tock" effect (Brochard et al., 2003). In this experimental design, the participants are presented with sequences of 13 to 16 identical tones in which 4 dB softer deviant tones inserted either at odd or even numbered positions. The participants are then asked to count the loudness changes in the sequences while their ERP responses are recorded. The results revealed ERP responses to deviants at odd positions in both groups, indicating an inherent trochaic grouping of non-linguistic sounds regardless of the native language of an individual.

In addition to grouping differences between the speakers of various languages, the studies investigating the ITL and the extent to which it impacts grouping biases reveal inconsistencies between the predictions of the ITL. Most of the studies reveal a consistent trochaic grouping preference for the intensity conditions but much less consistency for the duration conditions. Such a difference in grouping consistency indicate that some components of the auditory grouping biases are innate but some might be modified through linguistic experience. Growing evidence suggest the "nurture" component on top of a potential "nature" component, especially for the duration-based biases (Crowhurst, 2018). For instance, Yoshida et al. (2010) have reported that whereas 5-6-month old Japanese and English infants show no groupings for duration alternations, the groupings of 7-8-month old infants differ reflecting the phrasal structure of their native languages. Different influences of intensity and duration manipulations were reported in another study investigating infants' and adults' use of these cues in segmentation of tone sequences. Trainor and Adams (2000) presented both adults and 8-month-old infants sequences of tones where either duration or intensity of every third tone was manipulated. A gap detection paradigm was used, in which lower detection rates of silent gaps inserted at different locations indicated a group boundary. Whereas no group boundaries were revealed in the intensity condition, the duration condition revealed short-long groupings for both infants and adults, compatible with the duration principle of the ITL. Bion et al. (2011) showed that among the syllable pairs alternating in either pitch or duration, 7-month-old Italian infants group syllables with pitch alternations into trochees while they fail to group syllables with duration alternation into either iambs or trochees. This finding not only supports the view that duration biases might be learned through language experience or perceptual maturation (Langus et al., 2016), but also points to a possibility of universality in grouping biases based on pitch. The evolutionary path for the rhythmic grouping bias based on pitch was explored through comparative studies, which supported the proposal that it might be indeed innate. De la Mora et al. (2013) have tested the predictions of the ITL in rats by training them to discriminate pitch-alternating sequences of tones from sequences randomly varying in pitch. In the test trials, rats revealed trochaic groupings for pitch. The same procedure was followed for duration-alternating sequences but no grouping bias for duration sequences was obtained. The results of this study suggest the early emergence of the trochaic grouping bias based on pitch which may rely on perceptual bias shared by humans and non-human animals and that duration-based perceptual bias might depend on language experience. The possibility of shared perceptual biases among humans and other species is supported by evidence of basic perceptual grouping abilities observed in multiple animal species such as rats, pigeons, zebra finches and budgerians (Kotz et al., 2018). Hay & Diehl (2007) supports this view by attributing intensity and pitch based grouping biases to hard-wired auditory mechanisms.

Together these results obtained from developmental studies and comparative research indicate that some components of perceptual grouping biases are innate and some learned through experience, in humans possibly by 7-8 months of age, as when linguistic phrasal grouping develops (Yoshida et al., 2010).

The studies reported above investigate the ITL effects by manipulating either intensity, duration or pitch singly. Pointing out that this limitation of these studies and that these features vary together in speech, Crowhurst & Teodocio Olivares (2014) have investigated the joint influence of duration and intensity on rhythmic grouping of speech. They have compared the grouping biases of speakers of English and Zapotec, using sequences of syllables [de] and [ge]. Syllables were manipulated such that they were either violating one of the principles of the ITL (i. e., trochaic intensity and duration pattern of presented syllables) or complying with both principles of the ITL (i. e., trochaic intensity pattern and iambic duration pattern). The results confirmed earlier findings that English speakers group intensity alternations trochaically and duration alternations iambically. Zapotec speakers, in turn, grouped the duration alternations trochaically, in contrast with the predictions of the ITL in the conditions where only duration was manipulated. Moreover, in the conditions where both intensity and duration was manipulated, grouping patterns of Zapotec speakers suggested that intensity is a stronger predictor of grouping than duration.

Crowhurst (2018) have added vowel phonation to the other parameters influencing rhythmic grouping other than duration, intensity or pitch. In her study, speakers of English and Native Mexican Spanish were instructed to choose between two bisyllabic nonwords among the sequences of syllables alternating in either vowel duration or creaky phonation or both. The results revealed a bigger effect of phonation than duration on iambic grouping of the syllables. Contrary to earlier findings, there was a lack of consistent grouping for English speakers in the condition where only duration of the syllables was manipulated, supporting the inference that inconsistent results are obtained across the studies investigating ITL.

Bhatara et al. (2016) reported that stimulus properties play a critical role in auditory grouping. They compared the grouping preferences of speakers of German and French with chimeras, (i. e., instrumental syllables) that vary in complexity and in presence of pauses. The grouping preferences replicated the previous results that speakers of German have a more consistent grouping bias than speakers of French. Another analysis comparing the effects of the stimulus properties revealed that low variability sound stimuli, as the ones used in Hay & Diehl (2007), led to more ITL-compatible groupings, suggesting that experience with acoustic cues in a more ecological context is necessary for developing a robust grouping preference.

Taken together, studies investigating the predictions of the ITL or its universality reveal inconsistent results. While some studies reveal groupings compatible with the predictions of the ITL regardless of linguistic experience, others reveal differences between the speakers of languages with varying prosodic properties. The extent to which auditory grouping is governed by the ITL is also controversial, given that some of the studies reveal grouping differences for only linguistic stimuli whereas other reveal grouping differences influenced by the native language for non-linguistic stimuli or for both.

1.2 Turkish

The prosodic properties of Turkish are controversial and debated over a hundred years (Kabak, 2016). Though there are multiple approaches regarding its stress structure and its underlying and surface representations, consensus is on that Turkish has word-final stress at the word level (Göksel & Kerslake, 2004; Kabak & Vogel, 2001; Langus et al., 2016) and phrase-initial stress at the phrasal level (Nespor et al., 2008). The characteristics of Turkish stress will be explained further below.

1.2.1 Word level stress

1.2.1.1 Regular (final) stress

Turkish has word-final stress which is commonly assumed to be resulting from a stress assignment rule which places primary stress on the final syllable of a word, regardless of the word's length or weights of its syllables (Sezer, 1981). The prosodic word (PWd), which can be easily determined in Turkish, given that it is the domain of many prosodic processes such as vowel harmony, is the domain of "stress" in Turkish, and it is composed of a root and all suffixes (Kabak & Vogel, 2001; Özçelik, 2017). As can be seen in (1), stress moves forward each time a suffix is attached to a stem.

(1)

a.	kalém	'pencil'
b.	kalemlík	'penholder'
c.	kalemliklér	'penholders'
d.	kalemliklerím	'my penholders'
e.	kalemliklerimíz	'our penholders'
f.	kalemliklerimizdén	'from our penholders'

1.2.1.2 Exceptional (non-final) stress

Turkish also has many examples of non-final stress, which is generally classified into two types: exceptional root stress and exceptional affixal stress.

Exceptional root stress is limited to certain place names, (2a) person names (2b) and some uninflected adverbs (2c) and conjunctions as well as words which are barrowed (2d,e) (Examples taken from Kabak & Vogel, 2001).

(2)

a.	Ánkara	‘Ankara’
b.	Bárbara	‘Barbara’
c.	fákat	‘but’
d.	akváryum	‘aquarium’
e.	négatif	‘negative’

Exceptional affixal stress surfaces in two distinct ways, one involving pre-stressing suffixes and other involving stressed suffixes (Özçelik, 2017).

A pre-stressing suffix places word stress on the immediately preceding syllable, regardless of its rhymal profile (Özçelik, 2017). The majority of pre-stressing suffixes in Turkish is monosyllabic (3b-d), but there are a few bisyllabic pre-stressing suffixes, as illustrated in (4) (Kabak & Vogel, 2001).

(3)

a. atla-dí	b. atla-dí-*da*	c. atlá-*ma*-dı
jump-PAST	jump-PAST-too	jump-NEG-PAST
“He jumped.”	“He jumped, too.”	“He didn’t jump.”

d. atlá-*ma*-dí-*da*

jump$_{\text{-NEG-PAST}}$-also

"He didn't jump, either."

(4)

a. sabáh-*leyin*	b. durmák-*sızın*
morning-during	stop(v)-without
"in the morning"	"without stopping"

Turkish also has a set of stressed suffixes, indicated below (5), which are always bisyllabic and always stressed on their first syllable (Özçelik, 2017).

(5)

a. sor-únca	b. sor-árak	c. sor-úyor
ask-when	ask-by	ask$_{\text{-PRES.CONT}}$
"when (s)he asks"	"by asking"	"(s)he is asking."

d. sor-úyor-du-lar

ask$_{\text{-P.C-PAST}}$-Pl

"They were asking."

1.2.2 Stress at higher levels

Stress at the level of the phonological phrase (PPh) in Turkish falls on the leftmost prosodic word (PWd) (Kabak & Vogel, 2001), as exemplified in italic in (6).

(6)

a. [*o* kız]$_{PPh}$ b. [*başarılı* öğrenci]$_{PPh}$ c. [*şişman* kedi]$_{PPh}$

that girl successful student fat cat

"that girl" "successful student" "fat cat"

At the I-level, greater prominence falls on the rightmost phonological phrase within the intonational phrase, which is its head (Özçelik, 2017). Combined with the PP-level stress, sentential stress in Turkish falls on the leftmost prosodic word inside the rightmost phonological phrase in the IP, as illustrated in (7a). However, when there is an exceptional stress driving suffix is present in the rightmost prosodic word, it attracts stress (Özçelik, 2017), as in (7b).

(7)

a. [[*şişman* kedi]$_{PPh}$ [*yemek* yedi]$_{PPh}$]$_{IP}$

fat cat food eat$_{\text{-PAST}}$

"The fat cat ate food."

b. [[*şişman* kedi]$_{PPh}$ [yemek *ye-me-di*]$_{PPh}$]$_{IP}$

fat cat food eat$_{\text{-NEG-PAST}}$

"The fat cat ate food."

1.2.3 Acoustic correlates of Turkish stress

Whereas the theoretical aspects of Turkish stress have drawn attention of many linguists and have been debated about for decades, its phonetic realizations and acoustic correlates have lately become a focus. The acoustic realizations of Turkish stress have been explored at different levels of linguistic hierarchy.

Word-level, final stress in Turkish is mainly correlated with F0 changes (Levi, 2005; Pycha, 2006). Levi (2005) have investigated the intensity, duration and F0 differences between accented and unaccented syllables in Turkish words which were either finally or non-finally stressed. The results revealed that F0 is the most reliable cue signaling stress, followed by intensity and duration. In fact, she reports that duration differences between the accented and unaccented syllables were not perceptible. The discriminant analysis she performed reveals that F0 is the best predictor of stress, predicting over 93% of the data. Supporting the previous findings, Pycha's (2006) study reveals that stress in Turkish is signaled primarily by F0. The perceptual correlates of Turkish stress mirror its acoustic correlates. In their ERP study investigating the perceptual correlates of Turkish word stress and their contribution to lexical access, Zora et al. (2016) revealed that segmentally identical words with different lexical stress (e. g., [bebék] 'baby' and [bébek] 'a district in Istanbul') were distinguished solely on their prosodic features of F0, spectral emphasis and duration. Among these perceptual correlates, they found F0 to be the most prominent one, which they also report to be lexically specified in Turkish.

Though the consensus is on that F0 is the most robust cue signaling Turkish word level stress, it is realized differently in final and non-final stress. F0, intensity and duration measures of stressed syllables reported in Levi's (2005) study were higher in final positions than in non-final positions. Pycha (2006) reports that whereas F0 differs significantly between stressed and unstressed syllables in both final and non-final positions, intensity of the syllables does not differ significantly, revealing an asymmetry between the acoustic realizations of stress in different positions.

Investigating the phrasal prominence realizations in languages with different word order, Nespor et al. (2008) reports that Turkish has a phrase-initial phrasal stress which is signaled through pitch and intensity. She and her colleagues compare Turkish and French, languages with same, word-final stress and different word orders with the former being a complement-head (OV) and the latter being a head-complement (VO) language. They report that in Turkish, first stressed syllable of the phonological phrase, inside its complement prosodic word, has higher pitch and intensity values than the other syllables, adding that these differences result from the word order of Turkish.

1.3 Present study

The aim of the present study is to contribute to the literature on the grouping biases and the ITL by exploring grouping biases of native speakers of Turkish. To my knowledge, only one study investigates the predictions of the ITL with Turkish (Langus et al., 2016). Given that the most striking contrast in grouping biases is obtained with the speakers of Japanese (Iversen et al., 2008; Yoshida et al., 2010) and given the considerable similarities between Turkish and Japanese (Snape et al., 2009) in terms of word order and rhythm class (Levi, 2005), and given that it stands in the middle of the C and %V plane (see Figure 1), Turkish seems to be a suitable candidate to shed light on which aspects of language may be more decisive in rhythmic grouping.

Iversen et al. (2008) argues that the grouping differences between English and Japanese speakers are due to the opposite word order between the languages. Molnar et al. (2016) finds grouping differences between the speakers of two languages with opposite word order, namely, Spanish and Basque, supporting this view. However, the number of studies investigating native language effects on grouping with a head-final language is too small to make a strong inference based on comparisons.

The complex characteristics of Turkish in terms of its stress patterns and morphosyntactic structure leads to various predictions on how it might be reflected in auditory grouping. While its word level stress might induce an iambic grouping, its PP-level stress would induce the opposite. These predictions are discussed in the discussion section.

This study will make a contribution to the literature on the native language effects on rhythmic grouping in several ways. Firstly, the only one study that explored Turkish (Langus et al., 2016) has a few drawbacks, which I plan to overcome in this study. They used sequences consisting of the syllables, [pa], [su], [tu], [ke], [ma], [vi], [bu], [go], [ne], [du]. Within these syllables, there are three syllabic and one bisyllabic words commonly used in Turkish: [bu] ‘this’, [su] ‘water’, [ne] ‘what’ and [mavi] ‘blue’. Moreover, they used a method different than the ones used in most of the studies investigating ITL. They presented the sequences and asked the participants whether they heard a specific bi-syllabic (non)word in the sequence. In the current study, we will directly ask participants which grouping they perceived in a given stimulus sequence in order to make the results more comparable to the ones obtained by other studies that used the same method with linguistic (Hay & Diehl, 2007; Bhatara et al., 2013; Crowhurst & Teodocio Olivares, 2014) and non-linguistic stimuli (Hay & Diehl, 2007; Iversen et al., 2008; Molnar et al., 2016; Bhatara et al., 2016). Secondly, the current study aims to overcome some drawbacks of the Iversen et al. (2008) study. In the first experiment of the current study (Experiment 1A), the same sound stimuli from Iversen et al. (2008) was used. The experiment in their study was conducted in a classroom setting, where all participants heard the sequences from speakers and chose their groupings on paper. Even though this method is more ecologically valid and was used in other studies (Crowhurst & Teodocio Olivares, 2014), it has poor control over the setting. Moreover, Iversen et al. (2008) presented each trial once to the participant groups, which is an inadequate number for strong inferences based on the binary data it reveals. Also, they did not present the tone sequences that alternate in pitch, which will be included in the current study.

1.4 Research questions

1. How do native speakers of Turkish group sequences of tones, alternating either in amplitude, duration or pitch?
2. How do native speakers of Turkish group sequences of syllables, alternating either in amplitude, duration or pitch?
3. Do these grouping preferences reflect the prosodic properties of Turkish?

The extent to which language experience affects auditory grouping biases were investigated by presenting native speakers of Turkish tone and syllable sequences. The motivation to present both non-linguistic and linguistic sequences was to explore the extent to which native language effects, if any, emerge. The first experiment (Experiment 1A) was a replication of Iversen et al.'s (2008) study using the same set of stimuli. Participants were presented with tone sequences and asked to indicate the grouping they perceived. In Experiment 1B, the tone sequences were re-generated, rearranged and modified using the same parameters of Experiment 1A. The sequences in Experiment 1B were presented to participants following the same procedure used in Experiment1A with an increased number of trials. In the last experiment (Experiment 2), sequences consisting of the syllables [gü] and [kı] were used while keeping the procedure and the stimulus properties such as manipulation ratio similar. Together, the results of the experiments were reported and the implications were discussed with regard to the rhythmic properties of Turkish at different levels of linguistic hierarchy.

CHAPTER 2

EXPERIMENT 1

1.5 Experiment 1A

1.5.1 Method

1.5.1.1 Participants

Participants were 65 Boğaziçi University undergraduate students. They received 0,5 credit for introductory courses in return for their participation. Twenty-one of the participants reported to have received musical training. They had an average of 5.05 years of musical training (*SD* = 2.96, *Median* = 5). 12 participants reported to have received musical training at least for 5 years. All of the participants were native speakers of Turkish. On a 7-point Likert scale, participants were asked to indicate their level of English speaking and English comprehension. The average speaking level was 4.42 (*SD* = 1.21), and the average comprehension level was 5.75 (*SD* = .94). The participants reported to have started English education at an average age of 12.38 (*SD* = 4.16) until an average age of 19.88 (*SD* = 2.34).

1.5.1.2 Materials

Stimuli were 10-second sequences that consisted of tones which alternated in either amplitude, duration or pitch. The base tone was a 500 Hz complex tone (15 ms rise/fall) with a duration of either 150 or 250 ms. The tone was a low-pass square wave consisting of the fundamental and first three odd harmonics. The alternating tone was generated by multiplying either the amplitude, duration or the pitch of the base tone with the ratios of 1.5 or 2 for amplitude sequences; 1.25, 1.75 or 3 for duration sequences and 1.25, 1.5 or 2 for pitch sequences. A 20 ms gap was inserted between the base and the alternating tone. To eliminate possible effects of the starting tone, stimuli were faded in and out with durations of 2.5 seconds by applying a double-logarithmic ramp. Each sequence was generated both forward and reversed. These manipulations yielded a total of 32 sequences [2 base durations (150 or 250 ms) x 2 orders of presentation (forward or reversed) x 8 ratio parameters (2 amplitude ratios + 3 duration ratios + 3 pitch ratios)]. Six additional sequences were generated to be presented in the instruction phase, using different values of intensity, duration and pitch and different manipulation ratios than the ones used in the experimental phase. The participants were asked to fill out a questionnaire regarding their linguistic and musical experience, which included Goldsmiths Musical Sophistication Index (Gold-MSI) (Müllensiefen et al., 2014) (see Appendix A).

1.5.1.3 Apparatus

The sound stimuli were generated and manipulated in Matlab at CD quality 44.1 kHz sample rate using the Iversen et al. (2008) code. The experiment was prepared and run in PsychoPy.

The participants were presented the sound stimuli via Philips Shp 1900 stereo headphones.

1.5.1.4 Procedure

The experiment was conducted in dimly lit rooms at the Cognitive Processes Laboratory in Boğaziçi University. Firstly, participants received instructions on grouping choices. Keyboard responses were explained while presenting sequences with ratios other than the ones used in the experimental phase. The sequence order used in the instructions (e. g., short-long on the left, and long-short on the right of the screen) and the key locations were counterbalanced across participants. In the experimental phase, participants heard the sound sequence and pressed the left or right key to mark which of the two grouping options reflected the way they grouped them. They then rated how confident they were about their choice on a 3-point scale by choosing either “hiç emin değilim” (not sure at all), “biraz eminim” (somewhat sure) or “çok eminim” (very sure). After the experimental session, they filled out a questionnaire regarding their linguistic and musical experience and rated their music preference on a 7-level scale for Eastern to Western music. The experimental session lasted about 15 minutes and participants were tested individually.

1.5.1.5 Design

For duration manipulations, the overall design was a three-factor 3 (ratio: 1.25, 1.75 or 3) x 2 (base tone duration: 150 ms or 250 ms) x 2 (order of presentation: forward or reversed) within-participant design. For pitch manipulations, the overall design was a three-factor 3 (ratio: 1.25, 1.5 or 2) x 2 (base tone duration: 150 ms or 250 ms) x 2 (order of presentation: forward or reversed) within-participant design. For amplitude manipulations, the overall design was a three-factor 2 (ratio: 1.5 or 2) x 2 (base tone duration: 150 ms or 250 ms) x 2 (order of presentation: forward or reversed) within-participant design. The dependent variable was the proportion of iambic responses over all responses rated as “somewhat sure” or “very sure”.

1.5.1.6 Data Analysis

The responses were binary, with iambic responses coded as 1 and the trochaic responses coded as 0. The raw data obtained from PsychoPy[1] was imported to R[2]. The proportion of iambic responses over all responses were calculated for each condition. The same procedure was followed for (i) all responses, (ii) responses with confidence ratings of 2 ("slightly certain") and 3 ("completely certain") and (iii) responses with confidence rating of 3 ("completely certain") only. All following analyses will be based on responses with 2 and 3 confidence levels (see Table B1 in Appendix B for other results) because, unlike Iversen et al. (2008) who only analyzed responses with a confidence of 3, we believed there was no good reason to discredit responses with a confidence of 2.

The distributions of individual listeners' grouping preferences were computed and visualized. In order to reveal whether the proportion of iambic responses per collapsed condition differed from chance level (0.5), one-sample t-tests were run since this is the most widely used practice. Since the number of trials for each condition was insufficient to run parametric tests, individual conditions were pooled across manipulation.

1.5.2 Results

The proportions of each confidence rating for each manipulation condition are presented in Table 1.

[1] www.psychopy.org
[2] www.rstudio.com

Table 1. Percentages of Each Confidence Rating in Each Manipulation Condition.

		Manipulation			
		Duration	Amplitude	Pitch	*Total*
Confidence rating	"not sure at all"	9%	16%	13%	*13%*
	"somewhat sure"	35%	48%	46%	*43%*
	"very sure"	56%	37%	41%	*45%*

1.5.2.1 Distributions of grouping preferences

For Turkish listeners, there is a tendency to group the duration sequences iambically, with nearly half of the Turkish listeners (48%) choosing short-long grouping in more than 75% of the sequences. The grouping of amplitude and pitch sequences are more varied. 40% of Turkish listeners preferred a loud-soft (trochaic) grouping in amplitude sequences. There is no strong preference in pitch sequences: most of the listeners (35%) grouped the sequences as trochaically in half of the trials and iambically in the other half of the trials. The distributions of individual listeners' grouping preferences are presented in Figure 2.

1.5.2.2 One-sample t-tests

In order to assess whether grouping preferences for each manipulation differ from chance level (i.e., 0.5), one-sample t-tests were run. The results revealed significant differences from chance level only in duration condition. The sequences were grouped more iambically (M = .69, SE = .02, $t(64)$ = 8.35, p = .000). The average proportions of iambic responses for each manipulation are presented in Figure 3.

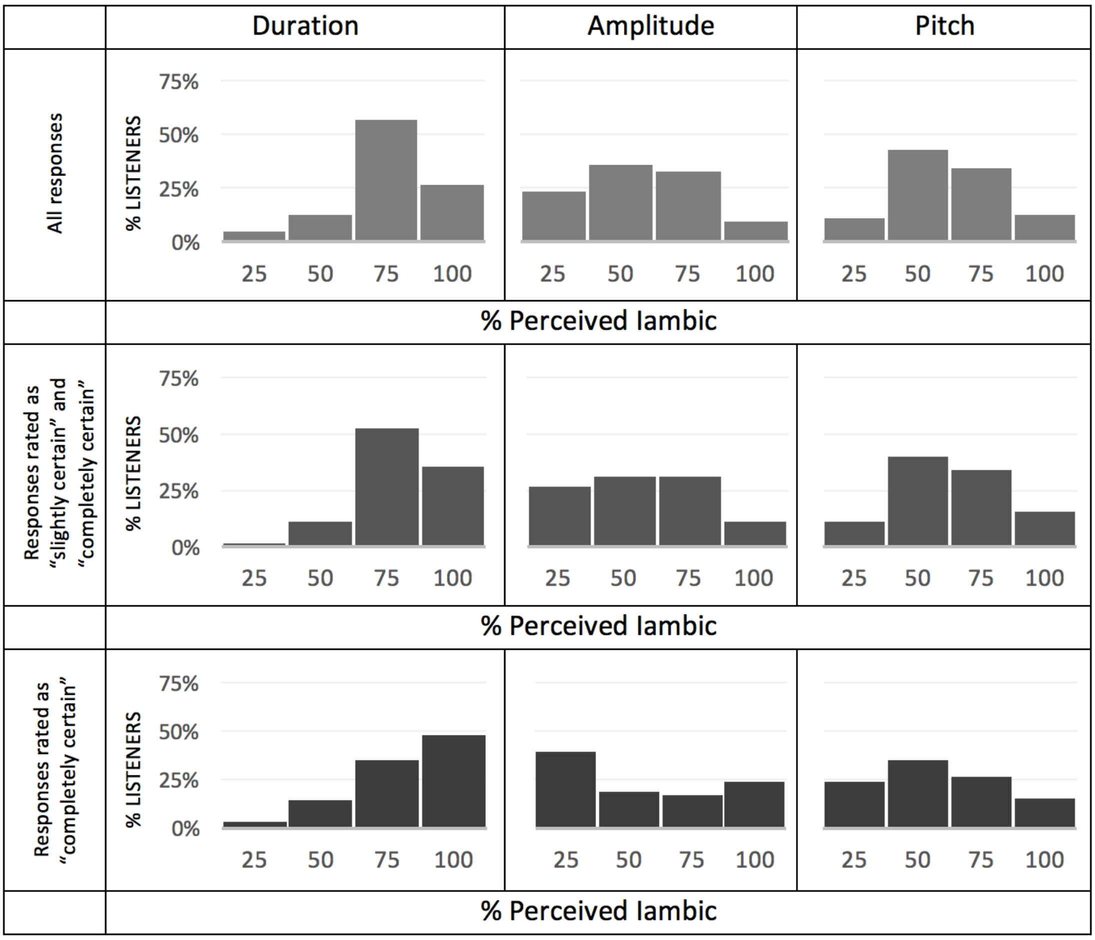

Figure 2 Distributions of individual listeners' grouping preferences. The columns represent duration, amplitude, and pitch manipulations, and the rows represent different levels of analysis. Graphs represent the distributions of all responses, responses with the confidence levels of 2 and 3, and responses with the confidence level of 3 only, in the first, second and third rows, respectively.

1.6 Experiment 1B

1.6.1 Method

1.6.1.1 Participants

Participants were 47 Boğaziçi University undergraduate students. They received 0,5 credit for introductory courses in return for their participation. Twelve participants reported experience with a musical instrument. Three participants reported to have received musical training at least for 5 years. All of the participants were native speakers of Turkish. On a 7-point Likert scale, participants were asked to indicate their level of English speaking and English comprehension. The average speaking level was 4.55 (*SD* = 1.19), and the average comprehension level was 5.86 (*SD* = .85). The participants reported to have started English education at an average age of 10.23 (*SD* = 2.99).

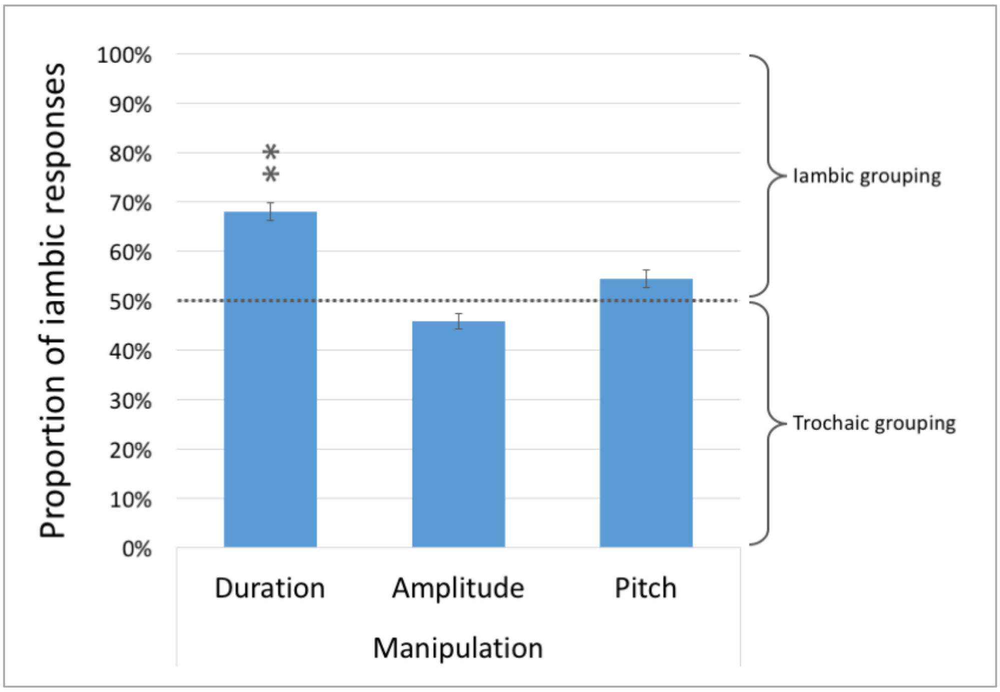

Figure 3 The average proportions of iambic responses for each manipulation. Error bars represent ±1 standard error of the mean (* $p < .001$).

1.6.1.2 Materials

The stimuli consisted of 10-second tone sequences formed by consecutive base and alternating tones. The base tone was a 300 Hz square-wave complex tone consisting of the fundamental and the first three odd harmonics with a duration of either 150 or 250 ms. The alternating tone was constructed by multiplying either the amplitude, pitch or duration of the base tone. The amplitude ratios were 1.5 and 2, the duration ratios were 1.25, 1.75 and 3 and the pitch ratios were 1.25, 1.5 and 2. A 20ms silent gap was inserted between the base and the alternating tone. The sequences had a duration of 10 seconds and they faded in and out for 3 seconds using a double-logarithmic ramp. In order to eliminate the effects of the starting tone, half of the sequences started with the base tone and the other half started with the alternating tone. The manipulations and 5 repetitions for each trial yielded 160 trials [8 (2 amplitude ratios + 3 duration ratios + 3 pitch ratios) x 2 (the starting tone) x 2 (base tone duration) x 5 (trials for each stimuli)]. Six additional sequences were generated to be presented in the instruction phase, using different values of intensity, duration and pitch and different manipulation ratios than the ones used in the experimental phase. The participants were asked to fill out a questionnaire regarding their linguistic and musical experience which included Goldsmiths Musical Sophistication Index (Gold-MSI) (Müllensiefen et al., 2014) (see Appendix A).

1.6.1.3 Apparatus

The sound stimuli were prepared and manipulated in Matlab. The experiment was constructed and presented in Matlab's Psychtoolbox extension (Kleiner et al. 2007).

1.6.1.4 Procedure

The experiment was prepared in Psychtoolbox extension of Matlab. The participants were seated in cubicles in a dimly lighted room. Participants were presented first with the instruction phase in which the grouping choices and the keyboard responses were explained, using sequences with ratios other than the ones used in the experimental phase. After completing half of the trials, the participants were offered a 10-minute break. In each trial, they were presented a sequence, and asked to press one of the buttons to choose one of the groupings and then to specify how confident they are of this choice by pressing one of the three buttons corresponding to a scale presented on the screen. Lastly, they were asked to fill out a questionnaire regarding their language education and musical background (see Appendix A). The experimental session is expected to last for 50 minutes. The position of the keys and the presentation order of the stimuli in the instructions were counterbalanced across participants.

1.6.1.5 Design

For duration manipulations, the overall design was a three-factor 3 (ratio: 1.25, 1.75 or 3) x 2 (base tone duration: 150 ms or 250 ms) x 2 (order of presentation: forward or reversed) within-participant design. For pitch manipulations, the overall design was a three-factor 3 (ratio: 1.25, 1.5 or 2) x 2 (base tone duration: 150 ms or 250 ms) x 2 (order of presentation: forward or reversed) within-participant design. For amplitude manipulations, the overall design was a three-factor 2 (ratio: 1.5 or 2) x 2 (base tone duration: 150 ms or 250 ms) x 2 (order of presentation: forward or reversed) within-participant design. The dependent variable was the proportion of iambic responses over all responses rated as “somewhat sure” or “very sure”.

1.6.1.6 Data Analysis

The responses were binary, with iambic responses coded as 1 and the trochaic responses coded as 0. The raw data obtained from MATLAB were imported to R. The proportion of iambic responses over all responses were calculated for each condition. The same procedure was followed for (i) all responses, (ii) responses with the confidence ratings of 2 ("slightly certain") and 3 ("completely certain") and for (iii) the responses with the confidence rating of 3 ("completely certain") only. All following analyses will be based on responses with 2 and 3 confidence levels (see Table B2 in Appendix B for other results).

In order to reveal whether the proportion of iambic responses over all responses in each condition differed from chance (0.5), one-sample t-tests were run.

In order to reveal the main effects of the variables of manipulation ratio, base tone duration and presentation order and their interactions, repeated measures ANOVAs were run. For the *amplitude* manipulations, a 2x2x2 ANOVA explored the main effects of manipulation ratio (high-to-low amplitude ratio: 1.50 or 2), base tone duration (150 ms or 250 ms), and presentation order (forward or reversed) and their interactions. For the *duration* manipulations, a 3x2x2 ANOVA explored the main effects of manipulation ratio (1.25, 1.75 or 3), base tone duration (either 150 ms or 250 ms), and presentation order (forward or reversed) and their interactions. For the *pitch* manipulations, a 3x2x2 ANOVA explored the main effects of manipulation ratio (1.25, 1.50 or 2), base tone duration (150 ms or 250 ms), and presentation order (forward or reversed) and their interactions.

The participants who reported to have experience with a musical instrument were referred as 'Musician' and those who reported not having any experience with a musical instrument were referred as 'Nonmusicians'. In order to equalize the sample sizes of Musicians ($N = 12$) and Nonmusicians ($N = 32$), 12 Nonmusicians were randomly sampled using random sampling function of SPSS. Independent samples t-tests were run for each (collapsed) condition in order to reveal the grouping differences between the participant groups.

On a 7-point Likert scale on Western music at one end and Eastern music on the other, the participants were asked to indicate where their preferred music genre resides. The results showed a skewness towards Western music. Musical preference was collapsed into three groups to cover the entire range with more or less equal representation: (1) Eastern music preference (N = 12) (2) Mixed music preference (N = 15) and (3) Western music preference (N = 18). A one-way ANOVA was run for each condition in order to explore the potential effects of music preference on grouping preferences.

In order to reveal the possible effects of musical sophistication, as indicated by Goldsmiths Musical Sophistication Index (Gold-MSI) (Müllensiefen et al., 2014) scores, participants were divided into two groups: (1) Low-MSI group (N = 12) with MSI scores lower than the mean of all scores (M = 59.32, SD = 12.03) and (2) High-MSI group (N = 22) with MSI scores higher than the mean. Independent samples t-test were run to explore the potential effects of musical sophistication on grouping preferences.

The possible effects of English proficiency on grouping was explored by comparing the BUEPT[3] scores of the participants, as self-reported in the questionnaire. There were only 5 participants who reported to have a BUEPT score of A. Therefore, differences between the participants who had scores of B (N = 22) and C (N = 17) were analyzed by running an independent samples t-test. In order to reveal the possible effects of the onset of the English education, the participants were divided into two groups, one with the onset age of 10 or below, one with the onset age of 11 or higher. A one-way ANOVA was run in order to explore the potential effects of the onset of English education on grouping preferences.

[3] Bogazici University English Proficiency Test (http://www.yadyok.boun.edu.tr/buept)

Another series of independent samples t-tests were run for each condition in order to reveal the possible effects of key location.

1.6.2 Results

1.6.2.1 Confidence ratings

The proportions of each confidence rating for each manipulation condition were presented in Table 2.

Table 2. Percentages of Each Confidence Rating in Each Manipulation Condition.

		Manipulation			
		Duration	Amplitude	Pitch	*Total*
Confidence rating	"not sure at all"	10%	17%	14%	*14%*
	"somewhat sure"	26%	37%	35%	*33%*
	"very sure"	64%	46%	51%	*54%*

1.6.2.2 One-sample t-tests

1.6.2.2.1 Collapsed conditions

One sample t-tests revealed significant deviation from the chance level only in the duration manipulation ($t(47) = 5.99$, $p = .000$, $d = .87$, $\beta = 1$). Participants grouped the duration sequences iambically ($M = .68$, $SE = .03$). The average proportions of iambic responses for each manipulation is presented in Figure 4.

1.6.2.2.2 Duration manipulations

In the duration manipulation conditions, proportions of iambic responses differed from chance level in two conditions when the ratio was 1.25 and in all conditions when the ratio was 1.75 or 3. When the ratio was 1.25, the proportion of iambic responses exceeded chance level when base tone duration was 150 ms and the sequence was presented forward ($M = .62$, $SE = .05$, $t(45) = 2.30$, $p = .026$, $d = .34$, β

= .64); and when base tone duration was 250 ms and the sequence was presented forward (*M* = .59, *SE* = .04, *t*(46) = 2.15, *p* = .000, *d* = .31, *β* = .57), both yielding iambic groupings.

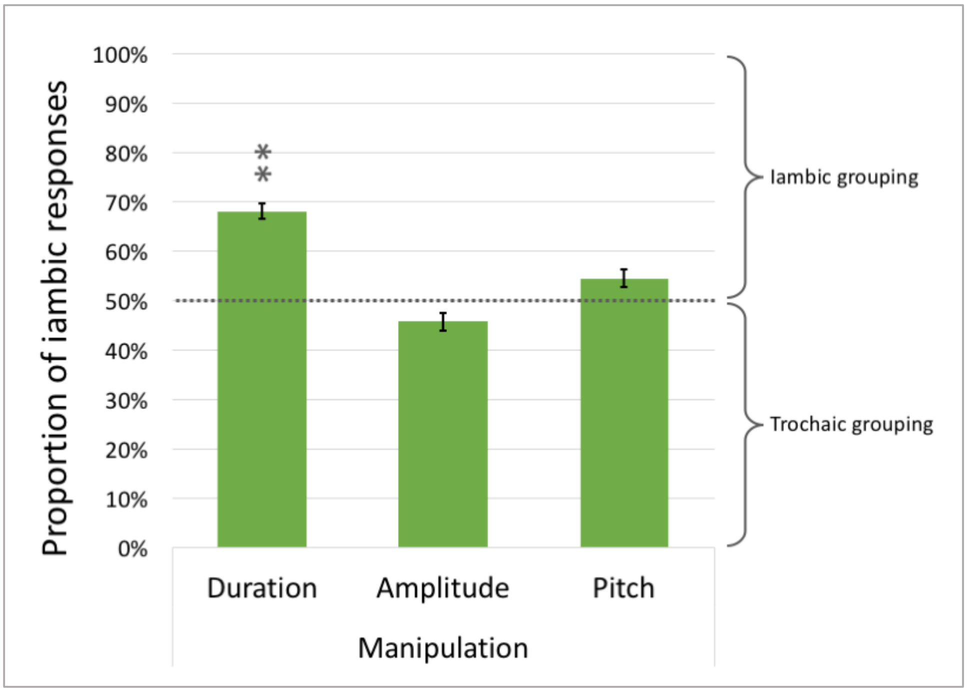

Figure 4 The average proportions of iambic responses for each manipulation. Error bars represent ±1 standard error of the mean (** *p* < .001).

When the ratio was 1.75 and base tone duration was 150 ms, proportion of iambic responses exceeded chance level both when the sequence was presented forward (*M* = .68, *SE* = .04, *t*(46) = 4.74, *p* = .000, *d* = .69, *β* = 1) and reversed (*M* = .67, *SE* = .04, *t*(46) = 3.88, *p* = .000, *d* = .57, *β* = .99). When the ratio was 1.75 and base tone duration was 250 ms, proportion of iambic responses exceeded chance level both when the sequence was presented forward (*M* = .74, *SE* = .04, *t*(46) = 5.86, *p* = .000, *d* = .85, *β* = 1) and reversed (*M* = .66, *SE* = .05, *t*(46) = 3.46, *p* = .001, *d* = .50, *β* = .91). When the ratio was 3 and base tone duration was 150 ms, proportion of iambic responses exceeded chance level both when the sequence was presented forward (*M*

= .78, *SE* = .04, *t*(46) = 7.12, *p* = .000, *d* = 1.04, *β* = 1) and reversed (*M* = .76, *SE* = .04, *t*(46) = 6.34, *p* = .000, *d* = .92, *β* = 1). When the ratio was 3 and base tone duration was 250 ms, proportion of iambic responses exceeded chance level both when the sequence was presented forward (*M* = .78, *SE* = .04, *t*(46) = 6.44, *p* = .000, *d* = .94, *β* = 1) and reversed (*M* = .70, *SE* = .04, *t*(46) = 4.57, *p* = .000, *d* = .67, *β* = 1), all yielding iambic groupings. The average proportions of iambic responses for each of the duration conditions are presented in Figure 5.

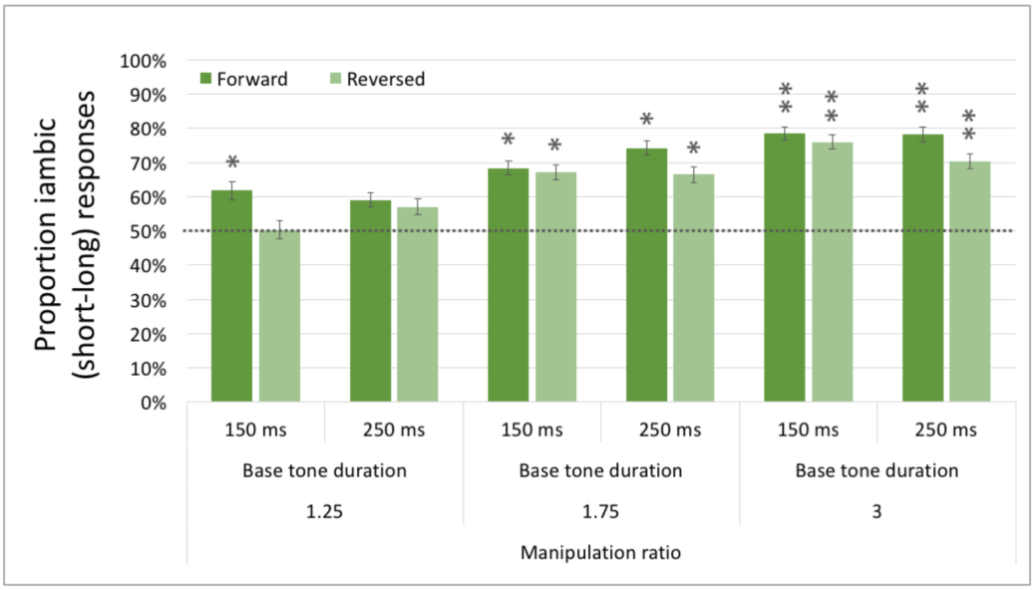

Figure 5 Average proportions of iambic responses for each of the duration conditions. Error bars represent ±1 standard error of the mean. (* $p < .05$, ** $p < .001$).

1.6.2.2.3 Amplitude manipulations

In order to reveal whether the proportions of iambic responses differ from chance level, one sample t-tests were run for each condition. The results revealed significant deviation from chance level only in one condition under amplitude manipulations. In the condition where manipulation ratio was 2, base tone duration was 250 ms and the sequence was presented forward, the average proportion of iambic responses (M = .39, SE = .05) were below chance level ($t(46) = -2.48$, $p = .017$, $d = .36$, $\beta = .68$), yielding a trochaic grouping. The average proportions of iambic responses for each of the amplitude conditions are presented in Figure 6.

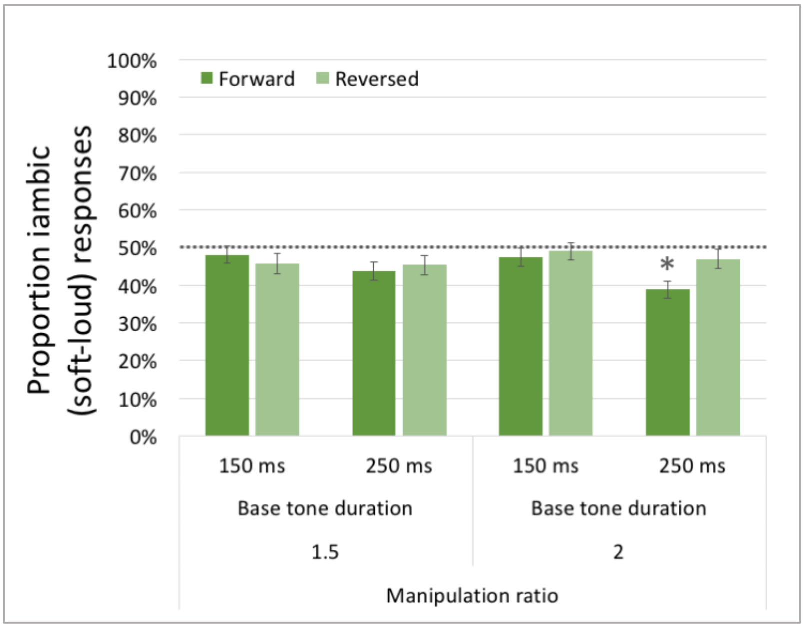

Figure 6 Average proportions of iambic responses for each of the amplitude conditions. Error bars represent ±1 standard error of the mean (* $p < .05$).

1.6.2.2.4 Pitch manipulations

In the pitch manipulation conditions, proportion of iambic responses differed from chance level ($t(45) = 2.24$, $p = .030$, $d = .33$, $\beta = .64$) only in the condition where the ratio was 1.25, base tone duration was 150 ms and the sequence was presented forward (M = .61, SE = .05), yielding an iambic grouping. The average proportions of iambic responses for each of the pitch conditions are presented in Figure 7.

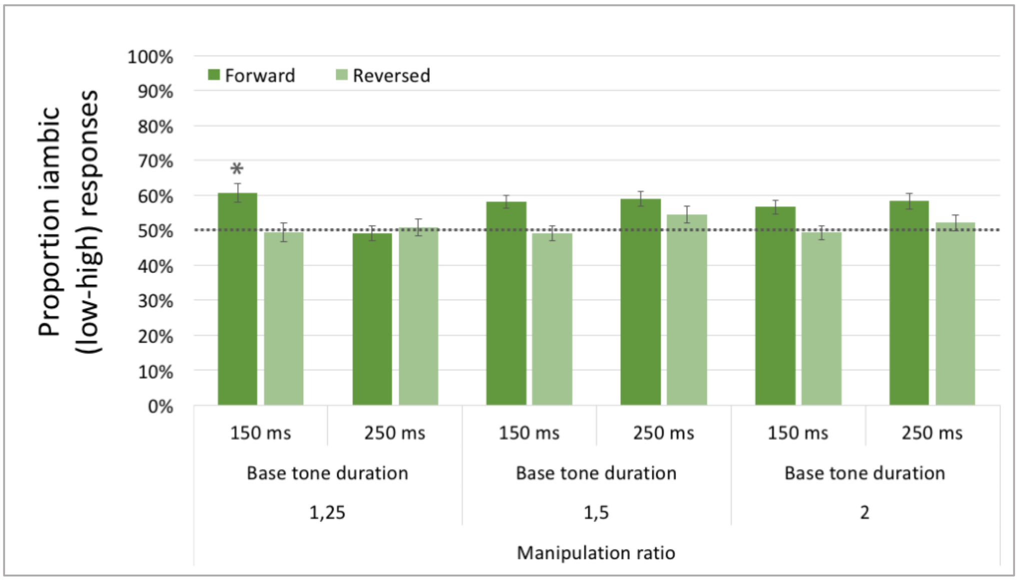

Figure 7 Average proportions of iambic responses for each of the pitch conditions. Error bars represent ±1 standard error of the mean (* $p < .05$).

1.6.2.3 ANOVAs

1.6.2.3.1 Duration manipulations

For duration manipulations, the main effect of ratio was significant [$F_{\text{Greenhouse-Geiser}}(1.62,72.91) = 19.97$, $p = .000$, $\eta_p^2 = .31$, $\beta = 1$]. Pairwise comparisons revealed significant differences between each of the manipulation ratios at the alpha level of .05. The sequences were grouped more iambically when the ratio was 3 ($M = .75$, $SE = .04$) than when it was 1.75 ($M = .69$, $SE = .03$), or when it was 1.25 ($M = .57$, $SE = .04$). The main effect of presentation order was also significant [$F(1,45) = 7.27$, $p = .010$, $\eta_p^2 = .14$, $\beta = .75$]. The sequences were grouped more iambically when they were presented forward ($M = .70$, $SE = .03$) than when they were presented reversed ($M = .64$, $SE = .03$).

1.6.2.3.2 Amplitude manipulations

For amplitude manipulations, there was a main effect of base tone duration [$F(1,45)$

= 4.88, *p* = .032, η_p^2= .10, *β* = .58]. The sequences were grouped more iambically when the base tone duration was 150 ms (*M* = .48, *SE* = .03) than when it was 250 ms (*M* = .43, *SE* = .03).

1.6.2.3.3 Pitch manipulations

For pitch manipulations, the main effect of presentation order was significant [*F*(1,45) = 8.60, *p* = .005, η_p^2 = .16, *β* = .82]. The sequences were grouped more iambically when they were presented forward (*M* = .58, *SE* = .04) than when they were presented reversed (*M* = .50, *SE* = .04).

1.6.2.4 The effects of music interest and experience

1.6.2.4.1 Preferred music genre

The results revealed a difference between the groups in a duration condition where the base tone duration of 150 ms was multiplied by the ratio of 1.75 and the sequence was presented reversed [*F*(2,43) = 3.40, *p* = .043, *MS* = .28, η_p^2= .14]. Western music listeners (*M* = .52, *SE* = .07) grouped this sequence more trochaically (long-short) than mixed music listeners (*M* = .76, *SE* = .07). Proportions of iambic responses did not differ between mixed music listeners (*M* = .76, *SE* = .07) and eastern music listeners (*M* = .75, *SE* = .09). The number of participants for each of the levels of the music preference scale is presented in Figure 8.

1.6.2.4.2 Experience with music

The tests revealed differences between Musicians and Nonmusicians. Pitch sequences were grouped more iambically by Musicians (*M* = .68, *SE* = .07) than Nonmusicians (*M* = .42, *SE* = .07) (*t*(22) = 2.75, *p* = .012, *d* = .98 , *β* = .65). In the amplitude condition where Musicians grouped the sequences more trochaically (M = .32, SE = .07) than Nonmusicians (M = .52, SE = .07), the difference between groups approached significance ($t(22) = -1.99$, $p = .059$, $d = .75$, $\beta = .47$). Average proportions of iambic responses of Musicians and Nonmusicians for duration, amplitude and pitch manipulations are presented in Figure 9.

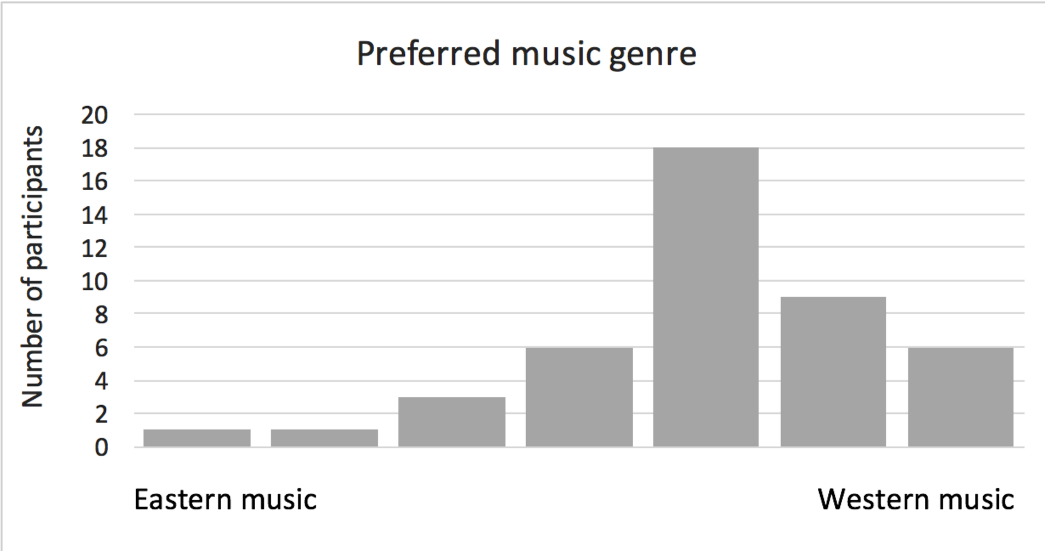

Figure 8 The number of participants for each of the levels of the music preference scale. The right end indicates Western music and the left end indicates Eastern music.

1.6.2.4.3 Gold-MSI

The results revealed no effect of musical sophistication as indicated in the Goldsmiths Musical Sophistication Index[4] (Gold-MSI).

[4] https://www.gold.ac.uk/music-mind-brain/gold-msi/

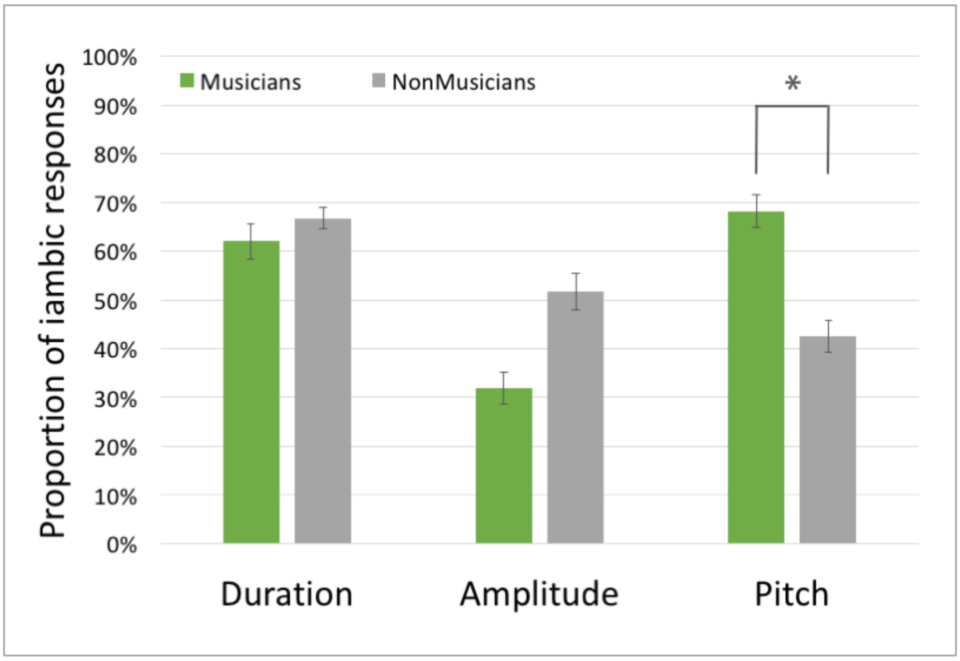

Figure 9 Average proportions of iambic responses of Musicians and Nonmusicians for duration, amplitude and pitch manipulations. Error bars represent ±1 standard error of the mean. (* $p < .05$)

1.6.2.5 The effects of second language proficiency

1.6.2.5.1 English proficiency

The results revealed differences between groups in a duration condition where the ratio was 1.75, base tone duration was 150 ms and the sequence was presented forward ($t(28.67) = 2.06$, $p = .048$, $d = .66$, $\beta = .49$). Participants with a BUEPT score of B (and thus higher proficiency in English) grouped these sequences even more iambically ($M = .77$, $SE = .04$) than the participants with a BUEPT score of C ($M = .60$, $SE = .07$). A significant difference was also found in a pitch condition ($t(37) = 2.46$, $p = .019$, $d = .75$, $\beta = .59$) where the ratio was 1.25, base tone duration was 150 ms and the sequence was presented forward. Participants with a BUEPT score of B (and thus higher proficiency in English) grouped these sequences more iambically (*M* = .62, *SE* = .07) than the participants with a BUEPT score of C (*M* = .38, *SE* = .06).

1.6.2.5.2 Onset age of English education

The results revealed no significant differences between the groups in overall duration [$F(2,43) = 1.82$, $p = .175$, $MSe = .08$, $\eta_p^2 = .08$], amplitude [$F(2,43) = .16$, $p = .854$, $MSe = .01$, $\eta_p^2 = .01$] or pitch [$F(2,43) = .19$, $p = .825$, $MSe = .01$, $\eta_p^2 = 0{,}009$] conditions

1.6.2.6 Control Variables

The results showed no effect of key location on grouping preferences in collapsed conditions.[5]

[5] There was a difference in a pitch condition where the manipulation ratio was 1.25, base tone duration was 250 ms and the sequence was presented forward ($t(45) = -2.37$, $p = .022$). The participants who were presented with the screen where the iambic grouping option was on the right (*N* = 24) showed more iambic groupings (*M* = .60, *SE* = .07) than the participants (*N* = 23) who were presented with the opposite screen arrangement (*M* = .38, *SE* = .06).

CHAPTER 3

EXPERIMENT 2

3.1 Method

3.1.1 Participants

Participants were 32 Boğaziçi University undergraduate students. They received 0,5 credit for introductory courses in return for their participation. Fifteen of the participants reported playing an instrument. One participant reported to have received musical training at least for 5 years. All of the participants were native speakers of Turkish. On a 7-point Likert scale, participants were asked to indicate their level of English speaking and English comprehension. The average speaking level was 4.73 (*SD* = 1.25), and the average comprehension level was 5.93 (*SD* = .91). The participants reported to have started English education at an average age of 9.67 (*SD* = 2.87).

3.1.2 Materials

The stimuli consisted of 10-second sound sequences formed by consecutive base and alternating syllables. The syllables were [gü] and [kı] and were uttered by a female native Turkish speaker. The recorded syllables were fixed at a fundamental frequency of 200 Hz. Then, their amplitude, duration and pitch values were manipulated. The amplitude ratios were 1.5 and 2, the duration ratios were 1.25, 1.75 and 3 and the pitch ratios were 1.25, 1.5 and 2. A 20ms silent gap was inserted between the base and the alternating syllables. The duration manipulation was applied only on the vowels (i.e, at the onset of vocalization). The sequences had a duration of 10 seconds and they faded in and out for 3 seconds using a double-logarithmic ramp. In order to eliminate the effects of the starting tone, half of the sequences started with the base syllable and the other half started with the alternating syllable. The manipulations and 5 repetitions for each trial yielded 160 trials [8 (2 amplitude ratios + 3 duration ratios + 3 pitch ratios) x 2 (the starting tone) x 2 (base tone duration) x 5 (trials for each stimuli)]. Six additional sequences were generated to be presented in the instruction phase, using the syllable [va], which was manipulated by different ratios than the ones used in the experimental phase. The participants were asked to fill out a questionnaire regarding their linguistic and musical experience, which included Goldsmiths Musical Sophistication Index (Gold-MSI) (Müllensiefen et al., 2014) (see Appendix A).

3.1.3 Apparatus

The syllables were recorded in Audacity, using Focusrite Scarlett Studio CM25 condenser microphone. Then, they were manipulated in Praat. The experiment was constructed and presented in Matlab's Psychtoolbox extension (Kleiner et al. 2007).

3.1.4 Procedure

The participants were seated in cubicles in a dimly lighted room. Participants were presented first with the instruction phase in which the grouping choices and the keyboard responses were explained, using sequences with ratios other than the ones used in the experimental phase. After completing half of the trials, the participants were offered a 10-minute break. In each trial, they were presented a sequence, and asked to press one of the buttons to choose one of the groupings and then to specify how confident they are of this choice by pressing one of the three buttons corresponding to a scale presented on the screen. Lastly, they were asked to fill out a questionnaire regarding their language education and musical background, which included Goldsmiths Musical Sophistication Index (Gold-MSI) (Müllensiefen et al., 2014) (see Appendix A). The experimental session is expected to last for 50 minutes. The position of the keys and the presentation order of the stimuli in the instructions were counterbalanced across participants.

3.1.5 Design

For duration manipulations, the overall design was a three-factor 3 (ratio: 1.25, 1.75 or 3) x 2 (syllable: [gü] or [kı]) x 2 (order of presentation: forward or reversed) within-participant design. For pitch manipulations, the overall design was a three-factor 3 (ratio: 1.25, 1.5 or 2) x 2 (syllable: [gü] or [kı]) x 2 (order of presentation: forward or reversed) within-participant design. For amplitude manipulations, the overall design was a three-factor 2 (ratio: 1.5 or 2) x 2 (syllable: [gü] or [kı]) x 2 (order of presentation: forward or reversed) within-participant design. The dependent variable was the proportion of iambic responses over all responses rated as "somewhat sure" or "very sure".

3.1.6 Data Analysis

The responses were binary, with iambic responses coded as 1 and the trochaic responses coded as 0. The raw data obtained from MATLAB were imported to R. The proportion of iambic responses over all responses were calculated for each condition. The same procedure was followed for (i) all responses, (ii) responses with the confidence ratings of 2 ("slightly certain") and 3 ("completely certain") and for (iii) the responses with the confidence rating of 3 ("completely certain") only. All following analyses will be based on responses with 2 and 3 confidence levels (see Table B3 in Appendix B for other results).

In order to reveal the main effects of the variables of the repeating syllable, manipulation ratio and presentation order and their interactions, repeated measures ANOVAs were run. For the amplitude manipulations, a 2x2x2 ANOVA explored the main effects of syllable (either [gü] or [kı]), manipulation ratio (either 1.50 or 2), and presentation order (forward or reversed) and their interactions. For the duration manipulations, a 3x2x2 ANOVA explored the main effects of syllable (either [gü] or [kı]), manipulation ratio (1.25, 1.75 or 3), and presentation order (forward or reversed) and their interactions. For the pitch manipulations, a 3x2x2 ANOVA explored the main effects of syllable (either [gü] or [kı]), manipulation ratio (1.25, 1.5 or 2), and presentation order (forward or reversed) and their interactions.

On a 7-point likert scale on Western music at one end and Eastern music on the other, the participants were asked to indicate where their preferred music genre resides. Since the number of participants in each level of the scale is inadequate for comparison, the scale is grouped into three levels: (1) Eastern music preference ($N = 11$) (2) Mixed music preference ($N = 10$) and (3) Western music preference ($N = 9$). A one-way ANOVA was run for each condition in order to explore the potential effects of music preference on the proportion of iambic responses.

The participants who reported to have experience with a musical instrument were referred as ‘Musician’ and those who reported not having any experience with a musical instrument were referred as ‘Nonmusicians’.

In order to reveal the possible effects of musical sophistication, as indicated by Goldsmiths Musical Sophistication Index (Gold-MSI) (Müllensiefen et al., 2014) scores, participants were divided into two groups: (1) Low-MSI group with MSI scores lower than the mean of all scores (*M* = 62.28, *SD* = 11.16) and (2) High-MSI group with MSI scores higher than the mean. Independent samples t-test were run to explore the potential effects of musical sophistication on grouping preferences.

The possible effects of English proficiency on grouping was explored by comparing the BUEPT scores of the participants, as self-reported in the questionnaire. There were only 5 participants who reported to have a BUEPT score of A, and 8 participants who reported to have a BUEPT score of B. Therefore, these participants are collapsed into a group with "high proficiency in English" (N = 13) and this group's grouping preferences were compared to the "low proficiency in English" group (N = 16) with the BUEPT scores of C, by running an independent samples t-test.

In order to reveal the possible effects of the onset age of English education, the participants were divided into three groups, one (Early-onset) with the onset age of 9 or below (*N* = 10), one (Mid-onset) with the onset age of 10 (*N* = 11) and one (Late-onset) with the onset age of 11 or higher (*N* = 9). Potential differences in grouping preferences between these groups were explored by running a one-way ANOVA.

In order to reveal the potential effects of control variables of key location and block order, separate series of independent samples t-tests were run.

3.2 Results

3.2.1 Confidence ratings

The proportions of each confidence rating for each manipulation condition were presented in Table 3.

Table 3. Percentages of Each Confidence Rating in Each Manipulation Condition.

		Manipulation			
		Duration	Amplitude	Pitch	*Total*
Confidence rating	"not sure at all"	8%	15%	15%	*13%*
	"somewhat sure"	28%	45%	43%	*39%*
	"very sure"	64%	40%	42%	*48%*

3.2.2 One-sample t-tests

3.2.2.1 Collapsed conditions

One sample t-tests revealed significant deviation from chance in grouping of duration ($t(31) = 7.91$, $p = .000$, $d = 1.40$, $\beta = 1$) and amplitude ($t(31) = -4.37$, $p = .000$, $d = .77$, $\beta = .99$) sequences. Participants grouped the amplitude sequences trochaically, given the average proportion of iambic responses ($M = .35$, $SE = .04$) was below chance. Duration sequences were grouped above chance ($M = .75$, $SE = .03$), yielding a iambic grouping. The average proportions of iambic responses for each manipulation are presented in Figure 10.

3.2.2.2 Duration manipulations

Under duration manipulations, proportions of iambic responses differed from chance level in all conditions when the ratio was 1.75 or 3. When the ratio was 1.75 and the repeating syllable was [gü], proportion of iambic responses was above chance both when the sequence was presented forward ($M = .83$, $SE = .04$, $t(31) = 7.92$, $p = .000$, $d = 1.40$, $\beta = 1$) and reversed ($M = .79$, $SE = .05$, $t(31) = 5.65$, $p = .000$, $d = 1$, $\beta = 1$). When the ratio was 1.75 and the repeating syllable was [kı], proportion of iambic responses was above chance level both when the sequence was presented forward ($M = .74$, $SE = .05$, $t(31) = 4.56$, $p = .000$, $d = .81$, $\beta = 1$) and reversed ($M = .79$, $SE = .05$, $t(31) = 6.44$, $p = .000$, $d = 1.14$, $\beta =1$).

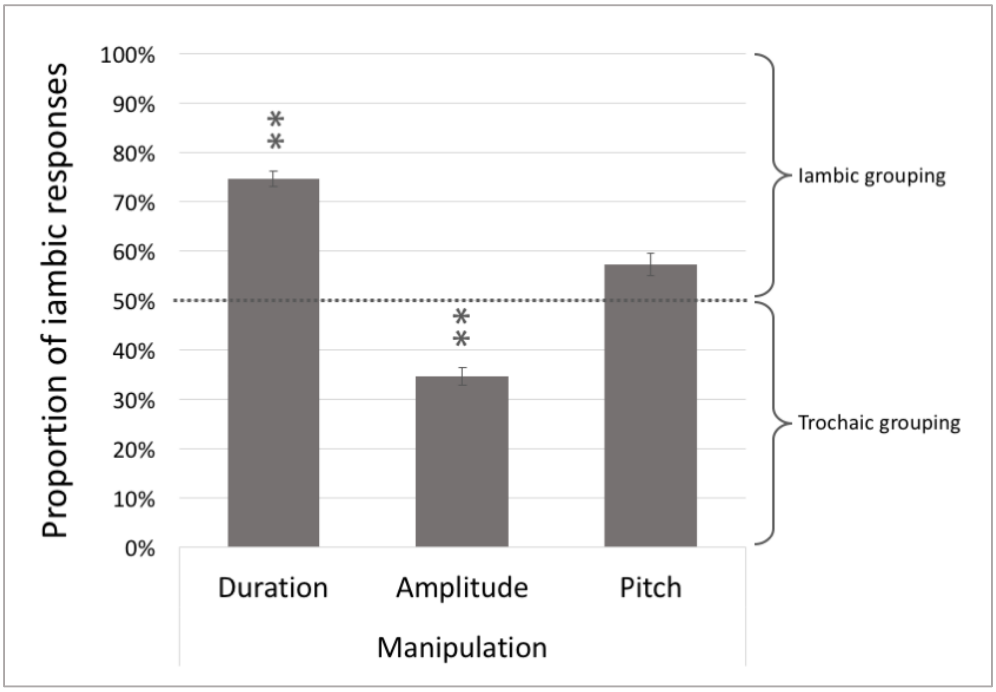

Figure 10 The average proportions of iambic responses for each manipulation. Error bars represent ±1 standard error of the mean (** $p < .001$).

When the ratio was 3 and the repeating syllable was [gü], proportion of iambic responses was above chance level both when the sequence was presented forward ($M = .93$, $SE = .03$, $t(31) = 14.64$, $p = .000$, $d = 2.59$, $\beta = 1$) and reversed ($M = .86$, $SE = .04$, $t(31) = 9.08$, $p = .000$, $d = 1.61$, $\beta = 1$). When the ratio was 3 and the repeating syllable was [kı], proportion of iambic responses exceeded chance level both when the sequence was presented forward ($M = .85$, $SE = .04$, $t(31) = 8.05$, $p = .000$, $d = 1.42$, $\beta = 1$) and reversed ($M = .84$, $SE = .04$, $t(31) = 8.37$, $p = .000$, $d = 1.48$, $\beta = 1$), all yielding iambic groupings. The average proportions of iambic responses for each of the duration conditions are presented in Figure 11.

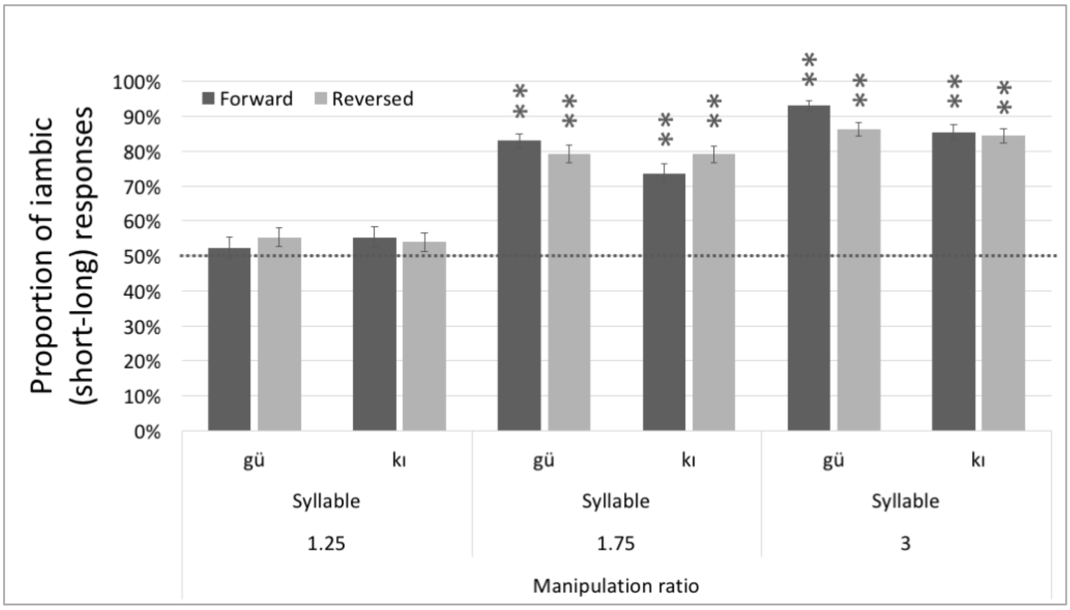

Figure 11 Average proportions of iambic responses for each of the duration conditions. Error bars represent ±1 standard error of the mean (** $p < .001$).

3.2.2.3 Amplitude manipulations

Under amplitude manipulations, results revealed significant deviations from chance level in the conditions where the ratio was 1.5 and the sequence was presented forward when the repeating syllable was [gü] ($M = .38$, $SE = .05$, $t(31) = 2.24$, $p = .032$, $d = .40$, $\beta = .65$). When the repeating syllable was [kı], the difference only approached significance ($M = .39$, $SE = .06$, $t(31) = -1.91$, $p = .066$, $d = .34$, $\beta = .52$). In conditions where the ratio was 2, proportions of iambic responses differed from chance level only when the sequences were presented reversed. These sequences were grouped trochaically both when the repeating syllable was [gü] ($M = .29$, $SE = .05$, $t(31) = -4.42$, $p = .000$, $d = .78$, $\beta = .99$) and when it was [kı] ($M = .25$, $SE = .05$, $t(31) = -5.10$, $p = .000$, $d = .90$, $\beta = 1$). The average proportions of iambic responses for each of the amplitude conditions are presented in Figure 12.

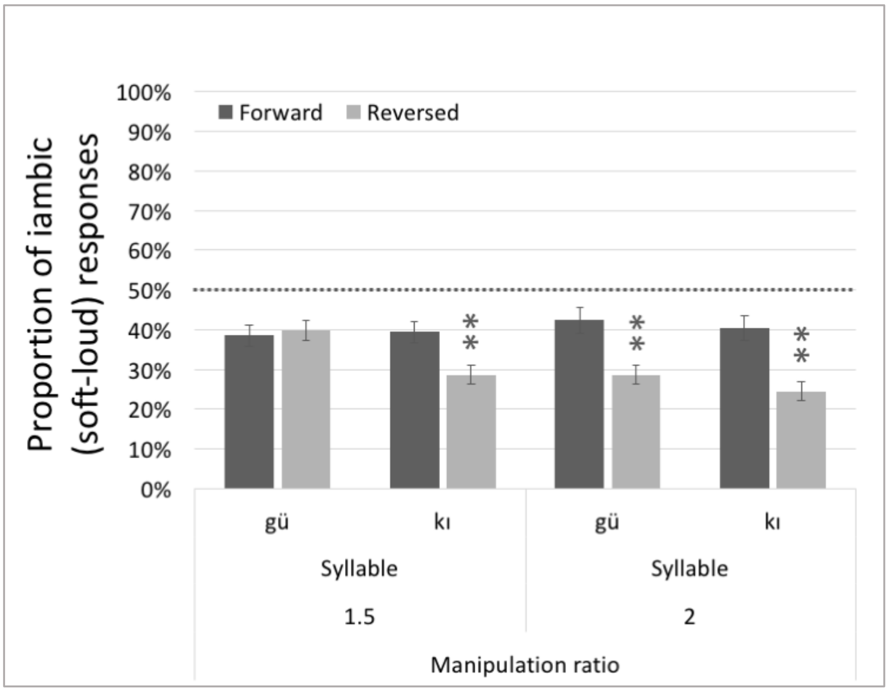

Figure 12 Average proportions of iambic responses for each of the amplitude conditions. Error bars represent ±1 standard error of the mean (** $p < .001$).

3.2.2.4 Pitch manipulations

Under pitch manipulations, results revealed significant deviations from chance level in the conditions where pitch of the [kı] syllable was multiplied with the ratio of 1.5 and the sequence was presented forward ($M = .62$, $SE = .06$, $t(31) = 2.23$, $p = .033$, $d = .40$, $\beta = .59$). Under these conditions, when the repeating syllable was [gü], the difference only approached significance ($M = .62$, $SE = .06$, $t(31) = 1.90$, $p = .067$, $d = .34$, $\beta = .49$) . In the condition where the pitch of [gü] syllable was multiplied by the ratio of 2, significant deviations from chance level was found when the sequence was presented reversed ($M = .63$, $SE = .05$, $t(31) = 2.53$, $p = .017$, $d = .45$, $\beta = .69$). In all these conditions, pitch sequences were grouped iambically. The average proportions of iambic responses for each of the pitch conditions are presented in Figure13.

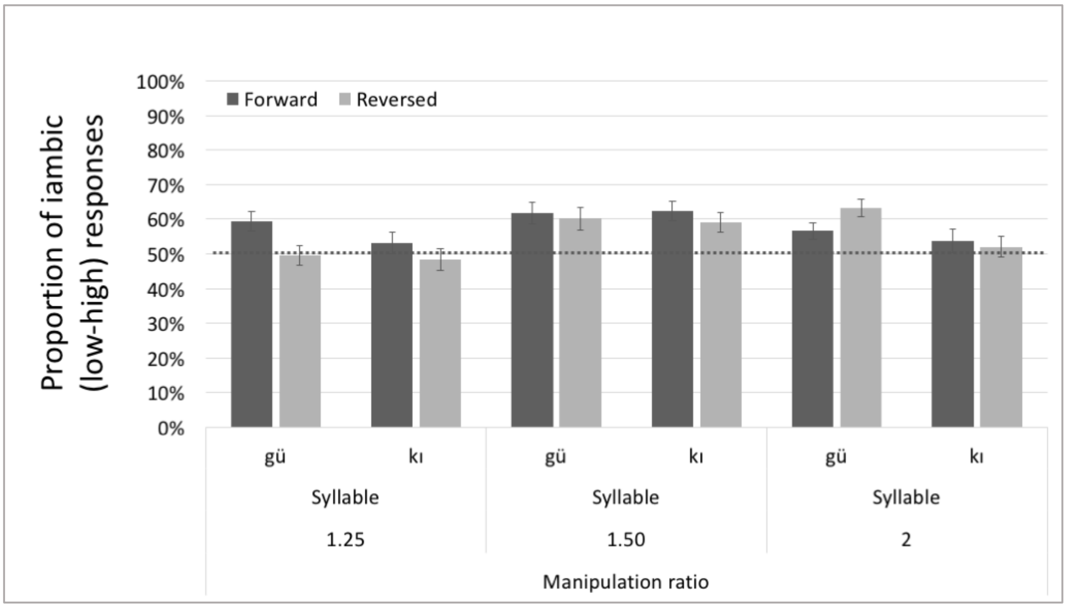

Figure 13 Average proportions of iambic responses for each of the pitch conditions. Error bars represent ±1 standard error of the mean (* $p < .05$, ** $p < .001$).

3.2.3 ANOVAs

3.2.3.1 Duration manipulations

The main effect of ratio was significant [$F_{Greenhouse\text{-}Geiser}(1.62,48.58) = 35.08$, $p = .000$, $\eta_p^2 = .54$, $\beta = 1$]. Pairwise comparisons revealed significant differences between each of the manipulation ratios at the alpha level of .05. The sequences were grouped more iambically when the ratio was 3 ($M = .87$, $SE = .05$) than when it was 1.75 ($M = .78$, $SE = .04$), or when it was 1.25 ($M = .54$, $SE = .05$).

3.2.3.2 Amplitude manipulations

3.3 There was a main effect of presentation order [$F(1,31) = 6.32$, $p = .017$, $\eta_p^2 = .17$, $\beta = .49$]. The sequences were grouped more iambically when the sequence was presented forward ($M = .40$, $SE = .04$) than when they were presented reversed ($M = .30$, $SE = .03$).

3.3.1.1 Pitch manipulations

The main effect of ratio [was significant $F(2,60) = 3.63$, $p = .032$, $\eta_p^2 = .11$, $\beta = .65$]. Pairwise comparisons revealed significant differences between the ratio of 1.25 and 1.50. Sequences with the ratio of 1.50 were grouped more iambically ($M = .61$, $SE = .05$) than sequences with the ratio of 1.25 ($M = .53$, $SE = .05$).

3.3.2 The effects of music interest and experience

3.3.2.1 Preferred music genre

The number of participants for each of the levels of the music preference scale is presented in Figure 14. The tests revealed no significant effect of music preference on grouping preferences.

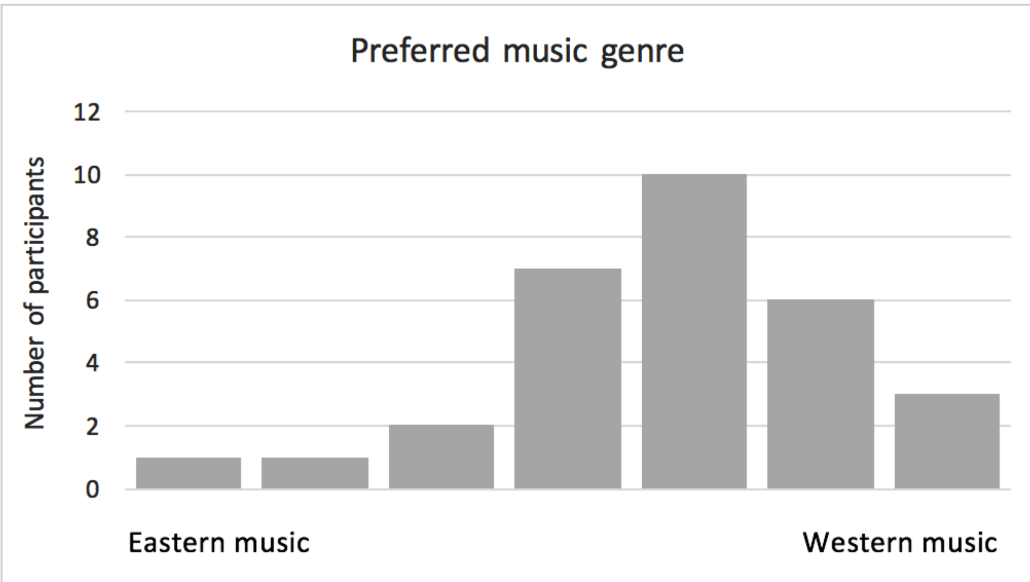

Figure 14 The number of participants for each of the levels of the music preference scale. The right end indicates Western music and the left end indicates Eastern music.

3.3.2.2 Experience with music

The tests revealed differences between Musicians (*N* = 15) and Nonmusicians (*N* = 15) in grouping of pitch sequences ($t(31) = -3.72$, $p = .001$, $d = 1.13$, $\beta = .87$). Pitch sequences were grouped more iambically by Musicians ($M = .73$, $SE = .05$) than Nonmusicians ($M = .46$, $SE = .05$). Average proportions of iambic responses of Musicians and Nonmusicians for amplitude, duration and pitch manipulations are presented in Figure 15.

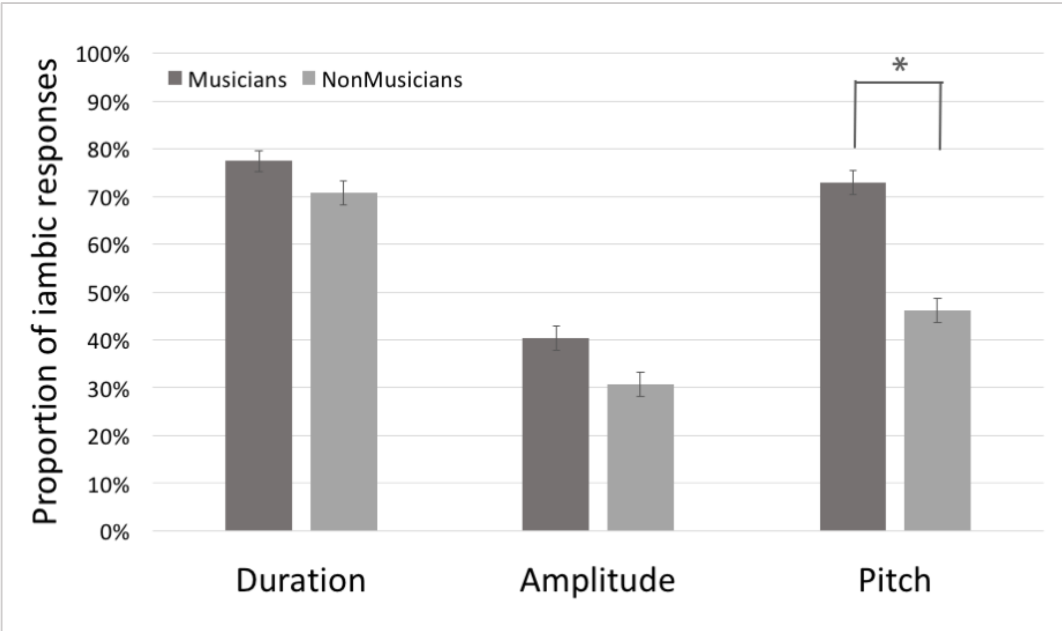

Figure 15 Average proportions of iambic responses of Musicians and Nonmusicians for amplitude, duration and pitch manipulations. Error bars represent ±1 standard error of the mean (* $p < .05$).

3.3.2.3 Gold-MSI

Independent samples t-tests revealed a difference between the groups in an amplitude condition where the repeating syllable was [gü], the ratio was 2 and the sequence was presented forward ($t(27) = -3.25$, $p = .003$, $d = 1.06$, $\beta = .73$). High-MSI group had a higher proportion of iambic responses (*M* = .65, *SE* = .10) than Low-MSI group (*M* = .23, *SE* = .06) in this condition. The same trend was observed in overall amplitude manipulations, but the difference between the High-MSI group (*M* = .42, *SE* = .06) and Low-MSI group (*M* = .29, *SE* = .04) only approached significance ($t(27) = -1.91$, $p = .067$, $d = 1.06$, $\beta = .39$).

3.3.3 The effects of second language proficiency

3.3.3.1 English proficiency

Independent samples t-tests revealed no significant differences between the "high proficiency in English" (*N* = 13) "low proficiency in English" (*N* = 16) groups in overall duration ($t(27) = .95$, $p = .349$, $d = .35$, $\beta = .15$), amplitude ($t(27) = 1.82$, $p = .964$, $d = .02$, $\beta = .03$) or pitch ($t(27) = -.82$, $p = .417$, $d = .30$, $\beta = .12$) conditions.

3.3.3.2 Onset age of English education

The results revealed differences between the groups in pitch condition [$F(2,19) = 4.15$, $p = .027$, $MSe = .19$, $\eta_p^2 = .31$]. Post-hoc tests using the Bonferroni correction with an overall alpha level of .05 revealed that the Early-onset group (*M* = .75, *SE* = .08) have grouped the sequences more iambically than Mid-onset (*M* = .51, *SE* = .05) and late-onset (*M* = .53, *SE* = .07) groups. The difference between the Mid-onset and Late-onset groups was not significant ($p = 1.00$).

3.3.4 Control Variables

3.3.4.1 Key location

An effect of key location was observed in overall duration manipulations, but the difference between the former group who were presented with the screen where the iambic grouping option was on the right (*M* = .80, *SE* = .03) and the latter group (*M* = .69, *SE* = .05) only approached significance ($t(30) = 1.82$, $p = .078$, $d = .62$, $\beta = .39$).[6]

3.3.4.2 Block order

The results revealed no difference between the groups of participants who were presented with opposite order of experimental blocks.

[6] In the specific conditions, an effect of key location was observed in a duration condition where the repeating syllable was [kɪ], manipulation ratio was 3, and the sequence was presented reversed ($t(20.28) = 2.58$, $p = .018$, $d = .83$, $\beta = .57$). The participants who were presented with the screen where the iambic grouping option was on the right (*N* = 15) showed more iambic groupings (M = .94, SE = .03) than the participants (*N* = 17) who were presented with the opposite screen arrangement (*M* = .76, *SE* = .07).

CHAPTER 4

DISCUSSION

4.1 Effects of manipulated variables

Three experiments were conducted in order to explore the rhythmic grouping preferences of native Turkish speakers. The first experiment (Experiment 1A) was a replication of the experiment in Iversen et al. (2008) study. Presented with sequences of tones alternating either in amplitude, duration or pitch, native Turkish speakers' grouping preferences were explored. The results revealed iambic (i.e., short – long) grouping of duration sequences and no strong preference in grouping of amplitude or pitch sequences. In the second experiment (Experiment 1B), sequences were re-generated, rearranged and were presented in 5 trials to another group of native Turkish speakers. The results were parallel to those of Experiment 1A: duration sequences were grouped iambically whereas no strong preference was found in grouping of amplitude or pitch sequences. In the final experiment (Experiment 2), the sequences consisted of the syllables [gü] and [kı] instead of square-wave tones, and were manipulated with the same parameters used in Experiment 1B. The results revealed trochaic (i.e., loud – soft) grouping of amplitude sequences and even stronger iambic grouping of duration sequences compared to the previous experiments. No strong preference was found in grouping of pitch sequences in Experiment 2.

Research questions:

1. How do native speakers of Turkish group sequences of tones alternating either in amplitude, duration or pitch?

2. How do native speakers of Turkish group sequences of syllables alternating either in amplitude, duration or pitch?
3. Do these grouping preferences reflect the prosodic properties of Turkish?

The results did not support the predictions regarding the grouping of duration sequences in Iversen et al. (2008) study, which they they stated as: “one can predict that native speakers of these languages (e.g., Turkish, Korean, and Marathi) will prefer long-short grouping” (p. 2270). Groupings of amplitude sequences were somewhat parallel to the results of Iversen et al. (2008) study. Surprisingly, the distributions of grouping preferences of Turkish speakers in the present study was similar to the distributions of grouping preferences of English speakers, rather than Japanese speakers in Iversen et al. (2008) study. The degree of similarity increased when the same analysis by selecting only the high-confidence responses was run on the present data. The results of their study and of the present study obtained with the same analysis procedure are presented in Figure 16. To my knowledge, only one study investigated native Turkish speakers’ rhythmic grouping of speech (Langus et al. 2017). In Langus et al.’s (2016) study, speakers of Italian, Persian and Turkish were presented with sequences of syllables and their sine-wave analogues that alternated either in duration or amplitude, and their grouping representations were explored. While language did not have an effect on grouping of the sine-wave analogues of syllable sequences, grouping patterns contrastive to ITL emerged with syllable sequences. Speakers of Persian and Turkish grouped the sequences trochaically (i.e., long-short) whereas Italian speakers grouped them iambically (i.e., short-long).

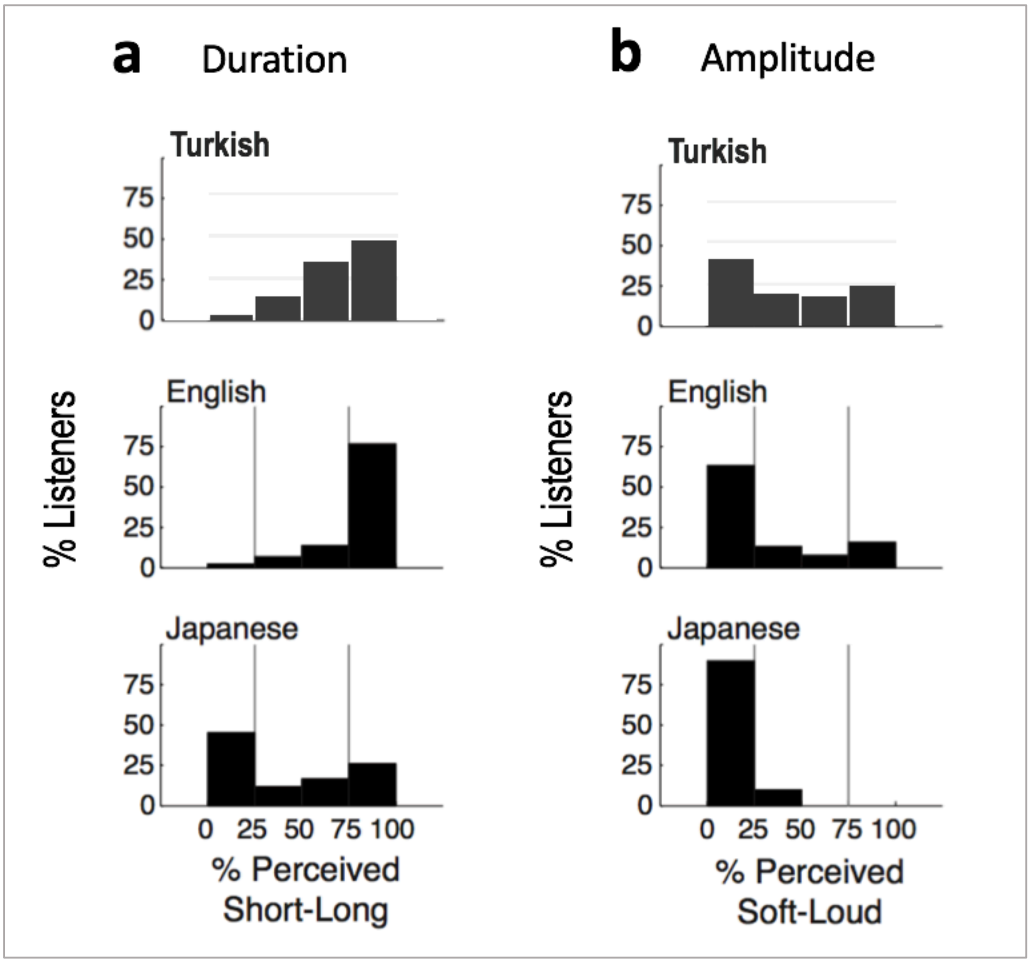

Figure 16 Comparison of the distributions of individual listeners' grouping preferences between the results of Iversen et al. (2008) and Experiment 1A (Figure is adapted from Iversen et al., 2008).

Results of the present study contradicts the results of Langus et al. (2016) study, as Turkish speakers in the present study grouped the duration sequences iambically, compatible with the ITL. This difference might be due to the drawbacks of their study or to the methodologic differences between the studies. They used sequences consisting of the syllables, [pa], [su], [tu], [ke], [ma], [vi], [bu], [go], [ne], [du]. Among these syllables, there are three syllabic and one bisyllabic word commonly used in Turkish, which may have adverse effects on the validity of the experiment and thus its results. Moreover, they used a method different than the one preferred in the present study as well as in the other studies investigating ITL, or in a typical ITL experiment (Hyde, 2011). On the other hand, the results of Langus et al.'s (2016) study show that the grouping representations and/or preferences of native Turkish speakers confirm ITL with stimuli which was insufficient to trigger linguistic cues. In the present study, sequences of tones alternating in duration were grouped iambically, as predicted by the ITL. Together these results show that even though there might be native language effects on grouping, it does not carry over to general auditory grouping for native Turkish speakers.

The results of the present study are incompatible with another study where differences in grouping between the speakers of languages with opposite word order were investigated (Molnar et al., 2016). The study revealed differences in grouping of tone sequences alternating in duration. Speakers of Basque, an OV language, grouped the sequences as long-short (trochaic) whereas speakers of Spanish, a VO language, grouped them as short-long (iambic), compatible with the ITL.

Given that the results do not support the prediction that speakers of OV languages would have similar grouping preferences which violate ITL, as observed in Iversen et al. (2008), Langus et al. (2016) and Molnar et al. (2016), these differences in grouping may not be deriving from the word order of the languages. This leads us to consider prominence at the lower levels of the linguistic hierarchy such as the level of intonational phrase, phonological phrase and prosodic word and how prominence is signaled and perceived at these levels.

ITL-compatible groupings found in the present study are in line with the findings of Hay and Diehl (2007). In their study, grouping preferences of English and French speakers were compared, using both linguistic and non-linguistic stimuli. Presented with sequences of either the repeating syllable [ga] or square wave segments, both English and French speakers grouped duration variations iambically and amplitude variations trochaically. The authors concluded from these results that ITL is a universal and domain-general mechanism governing rhythmic grouping. A closer examination of the properties of the languages at focus, however, leads to various inquiries regarding the validity of this conclusion. Similar results were obtained for French speakers in other studies investigating ITL. For instance, studies comparing grouping preferences of speakers of German and French have found that speakers of French had less consistent ITL groupings than speakers of German in grouping of syllable sequences (Bhatara et al., 2013) as well as non-linguistic sounds (Bhatara et al., 2016). In an ERP study where speakers of German and French learners of German were presented with German sentences with metrical and syntactic violations, Schmidt-Kassow et al. (2011) observed ERP responses to syntactic violations but not to metrical violations by French speakers, indicating lack of sensitivity to stress patterns for French speakers. This processing difficulty at perceiving stress patterns was referred to as stress 'deafness' (Dupoux et al., 1997) and was investigated in a number of studies with speakers of various languages such as European Portuguese (Lu et al., 2018), Dutch, Japanese, Persian and Indonesian (Rahmani et al., 2015); Finnish, Hungarian and Polish (Peperkamp & Dupoux, 2002). The common conclusion of these studies regarding the cause of this phenomenon is the predictable stress pattern of the language. In French, stress is not lexically marked, meaning that it does not carry lexical information. However, it predictably falls on the word's final vowel (Dupoux et al., 1997). This enables infants to deduce that stress is always word-final and does not need to be encoded (Peperkamp & Dupoux, 2002). During acquisition of a language with unpredictable stress, infants decide to keep stress information at phonological representations which, in turn, facilitates perception of stress later in life (Rahmani et al., 2015).

Though its reflections are observed in the findings of the studies which focus on ITL, perception of stress among speakers of various languages is most commonly tested by the Sequence Recall Task (SRT) (Dupoux et al., 2001). The goal of this task is to prevent access to the level of acoustic representation and hence to highlight the phonological level (Dupoux et al., 2008). Participants are presented with minimal pairs involving either a phonemic or a stress contrast. Asymmetrical results revealed by SRT between the perception of phonemic and stress contrasts indicate that "stress 'deafness'" is about storing prosodic features rather than perceiving them (Rahmani et al., 2015). For instance, Dupoux et al. (2008) have compared the performance of Spanish speakers and French late learners of Spanish and found that native French speakers, French learners of Spanish and Spanish speakers performed similarly at perceiving phonemic contrasts in a SRT experiment. However, performance at perceiving stress contrasts differed dramatically between Spanish and French speakers, indicating encoding difficulties of stress by French speakers. In the same study, performance differences between the speakers of Spanish and French in a lexical decision task demonstrated that stress deafness effect is not limited to encoding of stress to short-term memory and that it extends to lexical access.

The idea that a language exhibits noncontrastive stress leads its speakers to be 'deaf' to stress was supported by a number of studies. For instance, in Peperkamp & Dupoux (2002), SRT performances of speakers of Finnish revealed that they also exhibit stress 'deafness', indicating that this phenomenon is not limited to French and more importantly, is independent of the position of word stress, which is word-initial in Finnish, as opposed to French. The same study revealed stress deafness of speakers of Hungarian, another language with predictable word-initial stress.

Rahmani et al. (2015) have compared SRT performances of the speakers of Dutch, Japanese, French, Indonesian and Persian. They have suggested that lack of prosodic markings in the lexicon of a language would lead its speakers to exhibit processing difficulties in perceiving stress, regardless of the strength of the language's relation between stress location and word boundaries. The results revealed lower SRT performance and thus lower sensitivity to stress contrasts for the speakers of Persian, French and Indonesian but not for the speakers of Dutch and Japanese. That Persian has no transparent relation between accent and word boundary supported their suggestion that lack of lexically marked stress in a language is sufficient for its speakers to be stress-deaf. Lu et al. (2018) have reported processing difficulties for speakers of European Portuguese in absence of vowel reduction, indicating that stress 'deafness' may not be specific to languages with fixed stress.

Peperkamp & Dupoux (2002) have computed stress 'deafness' indexes for French, Finnish, Hungarian, Polish and Spanish by subtracting the mean percentage of errors made with the phonemic contrast from the mean percentage of errors made with the stress contrasts in SRT experiments. The results revealed differences in the magnitude of stress 'deafness' between the languages, indicating that this phenomenon reveals itself in different levels in different languages. This gradual exhibition of stress 'deafness' depends on how stress is signaled in a language and more specifically, whether it lexically marks stress.

Together with the consensus that Turkish has fixed stress (Kabak & Vogel, 2001; Göksel & Kerslake, 2004; Charette, 2008), results of the present study suggest that speakers of Turkish may also exhibit stress 'deafness'. In fact, Özçelik (2017) argues that Turkish has no stress, based on his proposal that it does not involve foot structure. His unified analysis of Turkish not only explains why speakers of Turkish may have decreased sensitivity to stress, it also reveals comprehensive solutions to problems with Turkish stress, which were studied by a number of linguists. He argues that the previous attempts to resolve these problems have not incorporated consideration of different acoustic correlates of stress in final and non-final positions and interaction of lexical stress at the phrasal level, leaving out many unaccounted cases. As stated in Chaper 1, Turkish has regular, word-final stress (Kabak & Vogel, 2001; Göksel & Kerslake, 2004). Regardless of the number of affixes attached to a word, it always falls on the final syllable of the word. Exceptional stress in Turkish is either rooted in some words, or it derives from affixation. Pre-stressing suffixes place stress on the immediately preceding syllable and are usually monosyllabic. Stressed suffixes are always bisyllabic and are stressed on their first syllable (Göksel & Kerslake, 2004).

According to Özçelik (2017), Turkish is a trochaic but footless language. Regular, final stress is a tone change that marks the word boundary, rather than lexically marked stress. Pre-stressing and stressed suffixes involve foot structure, as they come into the computation already footed in the input and thus are footed also in the output. When such an input foot is available, he argues, Turkish grammar can assign binary, weight-insensitive trochees. Otherwise, the grammar itself cannot parse syllables into feet. His formal analysis of feet in Turkish grammar accounts for the behavior of pre-stressing suffixes. Since the grammar ensures that feet are trochaic, binary and weight-insensitive, these footed suffixes end up stressing the preceding syllable at the surface representation. That stressed suffixes are always stressed on their first syllable is also explained by a trochaic, binary foot structure which places stress on the first syllable of its binary foot. His formal analysis, accounting also for exceptional root stress in Turkish, is summarized in Table 4.

Table 4. Summary of the Formal Analysis in Özçelik (2017). The examples and definitions are taken from Özçelik (2017). Feet are represented in brackets.

	Definition	Pre-stressing suffixes	Stressed suffixes	Exceptional root stress	
UR	Underlying Representation	/bekle-(me)-di/	/bak-(ınca)/	/an-(ka)-ra/	/(anka)-ra/
Align-Right	Align the right edge of a foot in the UR with the right edge of a foot in the SR	bekle(me)di	bak-(ınca)	an(ka)ra	(anka)ra
Ft-Bin	Foot binarity: <u>Yes</u> | No	bek(leme)di	bak-(ınca)	(anka)ra	(anka)ra
Trochaic	Foot shape: <u>Trochaic</u> | Iambic	bek(léme)di	bak-(ínca)	(ánka)ra	(ánka)ra
SR	Surface Representation	[bek(léme)di]	[ba(kínca)]	[(ánka)ra]	[(ánka)ra]

The different acoustic correlates of final and non-final stress in Turkish supports Özçelik's (2017) account. Whereas syllables with final stress exhibit only a slight F0 rise, non-final stress is correlated with both a sharp F0 rise and increased intensity (Levi, 2005; Pycha, 2006). It should be noted that some researchers found no strong correlate for final stress in Turkish (Konrot 1981, 1987, as cited in Özçelik, 2017). As Pycha (2006) states, "lack of clear acoustic correlates might suggest that final stress is nothing more than a percept for Turkish listeners" (p. 2)

If Turkish does not have lexically marked stress indeed, findings of the present study are not surprising. Under the assumption that Turkish grammar does not assign either iambic or trochaic foot, speakers of Turkish would have no representation of stress which would result in lack of strong preference for iambic or trochaic groupings in an ITL experiment, as found in the present study.

This line of thought would suggest that French may also be footless, given that speakers of French were reported to exhibit stress ‘deafness’ (Dupoux et al., 1997; Peperkamp & Dupoux, 2002; Rahmani et al., 2015) or lack of strong preference for grouping (Bhatara et al., 2013). This is exactly what Özçelik’s (2017) analysis of French reveals. He argues that French accent is phrase-final which is signaled through an optional, secondary high tone on the first or the second syllable of the first word. He supports his conclusion that French is footless, by reviewing a body of evidence which reveal that final prominence in French is a boundary tone rather than stress (Fery, 2001, as cited in Özçelik, 2017) and that the domain of prominence is the phonological phrase (PPh) in French rather than the word (PWd), which contradicts Hayes’s (1995) assumptions. He summarizes the arguments he makes in his study, as presented in Table 5.

Table 5. Summary of the Arguments Made in Özçelik (2017) Regarding the Foot Structure of Turkish and French. The table is taken from Özçelik (2017).

	Turkish	**French**
Does grammar have a means of parsing syllables into feet?	NO	NO
Does the language *ever* have feet?	YES (when pre-specified)	NO
How is prominence assigned (stress or intonation)?	Intonational for regular cases; stress when a foot is available	Intonational
At what level does intonational prominence apply?	PWd	PPh

Altmann (2006) brings a similar approach to the “stress ‘deafness’” phenomenon and classifies languages into a typology, as presented in Figure 17. Her series of experiments revealed that speakers of languages with predictable stress (i.e., Arabic, French, Turkish) had problems with locating stress whereas speakers of languages with no word-level stress (i.e., Chinese, Japanese, Korean) performed almost perfectly in the same task.

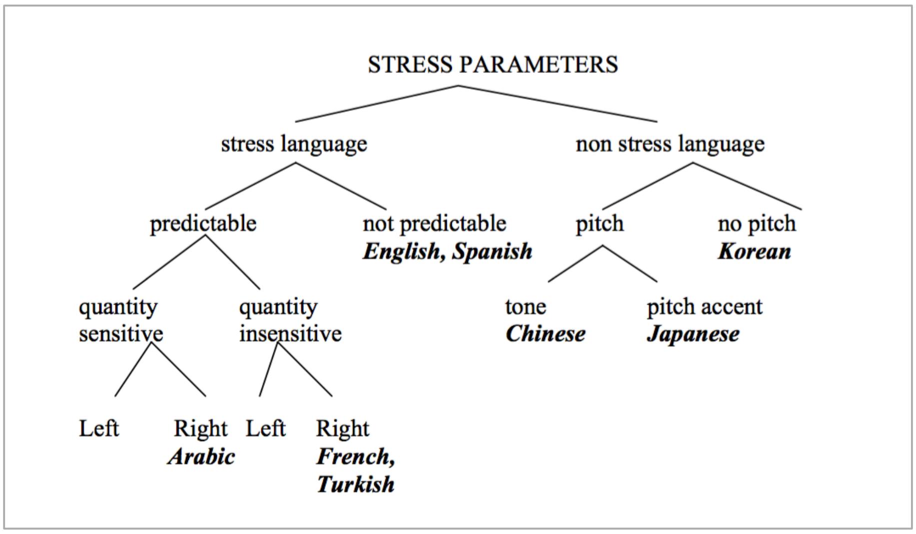

Figure 17 Stress typology model in Altmann's (2006) study. (Figure is taken from Altmann, 2006)

In contrast to the proposal that fixed stress in Turkish leads its speakers to have decreased sensitivity to stress, one may adduce the results of Langus et al.'s (2016) study where ITL-incompatible trochaic groupings of duration alternations were reported for speakers of Turkish and Persian. The evidence that speakers of Persian were reported to exhibit stress 'deafness' (Rahmani et al., 2015) speaks against the conclusions of Langus et al.'s (2016) study. Moreover, the study employs a method not used in any of the studies investigating ITL, which supports the idea that the findings of the study may be due to methodological factors. The present approach also justifies why the prediction that Turkish speakers would have grouping preferences similar to Japanese speakers (Iversen et al., 2008) was not supported by the present findings. Though Turkish and Japanese share common characteristics (Snape et al., 2009; Levi, 2005), Japanese differs from Turkish on the ground that it exhibits variable stress (Rahmani et al., 2015). That Japanese speakers outperform Turkish speakers in perception of stress (Altmann, 2006) supports this idea.

An unexpected result of the present study was lack of strong preference for grouping of amplitude sequences with tonal stimuli, both when the repeating square-wave tone was 500 Hz (Experiment 1A) and 300 Hz (Experiment 1B). If ITL is a domain-general mechanism, trochaic grouping of tones should have been observed in the results of these experiments. On the other hand, Bhatara et al. (2016) point out that stimulus with low variability, (as the stimulus in the present study is) may cause a pseudo-ceiling effect by placing no demand on the participants to process the stimuli at an abstract level.

4.2 Effects of musical experience

Bhatara et al. (2016) reports that musical experience has a facilitating effect on the consistency of grouping. In the study, speakers of French who received music education had more ITL-consistent groupings than those who have not received music education. Similar trend of results was obtained in the present study with Turkish speakers. Both in Experiment 1B with tone sequences and Experiment 2 with syllable sequences, participants who reported to play an instrument had more iambic groupings in pitch condition than those who reported to have no experience with music. In musical terms, this indicates that the former group perceived the tones as ascending. The are other instances of individual differences in pitch perception, such as “the tritone paradox” (Deutsch, 1991). In her studies, Deutsch (1991) found that among the same pair of tones with a half-octave interval (i. e., a tritone) participants with different linguistic backgrounds perceive opposite patterns. Together with the results of the present study, language-based effects on pitch perception indicate that drawing conclusions about the nature of pitch perception is difficult and requires further, careful experimentation.

A striking difference between the results of Experiment1B and Experiment 2 was the asymmetry between musicians and Nonmusicians in the extent of trochaic grouping of amplitude sequences. In the experiment with tone sequences, participants with musical experience had more trochaic groupings. However, in the experiment with syllable sequences, there was only a slight, insignificant difference between the groups with Nonmusicians having more trochaic groupings. These results are in line with those obtained in Bhatara et al.'s (2016) study in which musicians with French, but not German, as their native language grouped sequences of "chimaras"(i.e., musical syllables) more trochaically. Lack of linguistic cues in experiments where tones are presented may lead the participants to rely on musical representations rather than linguistic ones, especially when participants have no or weak representations. With a bigger sample, we expect the difference between Musicians and Nonmusicians observed in the results of the present study to appear significant.

Given the shared characteristics between music and language (Patel, 2003), the possible effects of music on rhythmic grouping cannot be overlooked. Experience with different musical styles may have interfered with the grouping preferences of the participants in the present study. On the other hand, attempts to rule out the effects of music on grouping is difficult given the circular structure of both rhythmic and tonal aspects of music. Though differences in grouping of musical rhythm can be obtained across cultures (Polak et al., 2018; Jacoby & McDermott, 2017), less can be concluded regarding the direction or order of the grouping representations.

4.3 Limitations

In the present study, participants had varying degree of proficiency in English. Since English is mandatory in Turkey's education system, native speakers with no proficiency in English could not be recruited. In a study -investigating grouping preferences of Spanish and Basque speakers (Molnar et al., 2016) there were four participant groups with varying degrees of proficiency in Spanish and Basque. Preference for either iambic or trochaic grouping was observed only for Spanish Monolinguals and Basque-dominant Bilinguals, but not for Spanish-dominant Bilinguals or Basque-proficient Bilinguals. These results indicate effects of proficiency in a second language with opposing word order on grouping. Even though rhythmic grouping biases are reported to emerge in infancy (Hay & Saffran, 2012; Toro & Nespor, 2012), the fact that participants were to some degree proficient in English may have deteriorated the present findings.

The present study employed a method where one can switch from one grouping to another in a trial. Even though most of the studies use this method to investigate ITL, another method such as clapping can be adopted to reveal grouping representations rather than preferences. Such an approach might also eliminate effects of key location as revealed in the present findings.

Significant effects of the presentation order were obtained in Experiment 1B for duration and pitch conditions and in Experiment 2 for amplitude condition. More iambic preferences were obtained in these conditions when the sequences were presented forward, namely, iambically. This drawback of the present method can be eliminated by increasing the fade duration and adding masks to the sequences. Bhatara et al. (2016) have masked the sequences used in the experiments for 3 seconds by white noise, fading out according to a raised-cosine function. The sequence itself was also faded in from silence. They report no effects of presentation order. A future study could use the same method to see whether presentation order still leads to differential effects.

4.4 For future research

The present study involves participant groups with varying degrees of proficiency in a second language. However, the sample sizes are small to make a comparison between the groups. In the follow-up studies, bigger samples involving participants with both high and low proficiency in second language should be recruited. Moreover, conclusions regarding the effects of one's native language on grouping would be best drawn by recruiting participants with different native languages.

The present study reveals grouping preferences of native speakers of Turkish in an ITL experiment. Though the results support the idea that Turkish is footless and thus its speakers exhibit stress 'deafness', Turkish speakers' sensitivity to stress contrasts can be best revealed by a SRT experiment. On the other hand, that Turkish has stress contrasts such as [bebék] 'baby' and [bébek] 'a district in Istanbul' may enable its speakers to perceive such contrasts in a SRT design.

Özçelik (2017) argues that if his arguments regarding the footless analysis of French and Turkish are correct, it would open up the possibility of reclassifying other languages which have fixed stress as footless. The languages whose speakers were reported to exhibit stress 'deafness', such as Finnish, Hungarian (Peperkamp & Dupoux, 2002); Persian, Indonesian (Rahmani et al., 2015); and European Portuguese (Lu et al., 2018) can be reanalyzed from Özçelik's (2017) perspective in future research.

4.5 Conclusion

Rhythmic grouping preferences of native Turkish speakers were explored by running three experiments. Results revealed no strong preference in grouping of tones alternating in either amplitude or pitch, and an iambic preference for grouping of tones alternating in duration. The following experiment showed that native Turkish speakers group sequences of syllables trochaically when they alternate in amplitude and iambically when they alternate in duration. The results are mostly compatible with the predictions of the ITL, except for the finding of an absence of (strong) grouping preference for amplitude sequences. The research question was whether Turkish as a native language influences rhythmic grouping of speech and non-speech stimuli. Though the results suggest no native language effects at first glance, in contrast to what was predicted (Iversen et al., 2008) or observed (Langus et al., 2016) by the previous studies, closer examination of the prosodic properties of Turkish suggest that there may indeed be native language influences but in a different way than was proposed by these.

The results suggest no native language effects under the assumption that Turkish has lexically marked stress, which would render grouping preferences of Turkish speakers as incompatible with the ITL. On the other hand, under the assumption that Turkish has fixed, predictable stress, ITL-compatible groupings are no surprise, given that the speakers of languages with fixed, predictable stress were reported to have decreased sensitivity to stress deviations, which was referred as "stress 'deafness'" (Dupoux et al., 1997; Peperkamp & Dupoux, 2002). A comprehensive phonological analysis of Turkish as footless (Özçelik, 2017) in addition to different acoustic correlates between final and non-final stress in Turkish (Levi, 2005; Pycha, 2006) supports this latter assumption. Moreover, results of the present study are parallel to those obtained for speakers of French in ITL studies (Schmidt-Kassow et al., 2011; Bhatara et al., 2013; Bhatara et al., 2016), who were reported to exhibit stress ‘deafness’ (Dupoux et al., 1997; Peperkamp & Dupoux, 2002). That French was also analyzed as a footless language (Özçelik, 2017) further supports the conclusion that speakers of Turkish may have stress ‘deafness’. However, the fact that Turkish also has stress contrasts suggest otherwise and renders it less amenable for an investigation of stress insensitivity in Turkish speakers. Further research is needed to investigate these proposals in depth. Taken together, the present study draws attention to the consideration of a wide range of linguistic properties of a language when formulating hypotheses of auditory grouping.

APPENDIX A

QUESTIONNAIRE

1. Katılımcı numarası *(Participant number)*
2. İşitme ile ilgili bir sorununuz var mı? *(Do you have a hearing problem?)*
    - Evet (lütfen deney yürütücüsünü bilgilendirin) *(Yes (please inform the experimenter))*
    - Hayır *(No)*
3. Yaşınız: *(Age)*
4. Cinsiyetiniz: (*Gender)*
    - Kadın *(Female)*
    - Erkek *(Male)*
    - Diğer *(Other)*
5. El tercihi: *(Hand preference)*
    - Sağlak *(Right handed)*
    - Solak *(Left handed)*
6. Anadiliniz: *(Your native language)*
    - Türkçe *(Turkish)*
    - Diğer *(Other)*
7. Konuştuğunuz diller *(Language spoken)*
    - Kürtçe *(Kurdish)*
    - İngilizce *(English)*
    - Almanca *(German)*
    - Fransızca *(French)*
    - İspanyolca *(Spanish)*

- İtalyanca *(Italian)*
- Diğer (yukarıda belirtilen) *(Other (as specified above))*

8. Konuştuğunuz diller ve bu dillerdeki seviyeniz *(Proficiency in the languages spoken)*

1(en düşük seviyede) 2 3 4 5 6 7 (en iyi seviyede)

1(lowest level) 2 3 4 5 6 7 (highest level)

- Kürtçe *(Kurdish)*
- İngilizce *(English)*
- Almanca *(German)*
- Fransızca (*French)*
- İspanyolca *(Spanish)*
- İtalyanca *(Italian)*
- Diğer (yukarıda belirtilen) *(Other (as specified above))*

9. İngilizce yeterlilik sınavı (BUEPT / Proficiency) sonucunuz: *(English proficiency test / BUEPT score)*
10. İngilizce konuşma seviyeniz *(English speaking level)*

    1(en düşük seviyede) 2 3 4 5 6 7 (en iyi seviyede)

    1(lowest level) 2 3 4 5 6 7 (highest level)
11. İngilizce anlama seviyeniz *(English comprehension level)*

    1(en düşük seviyede) 2 3 4 5 6 7 (en iyi seviyede)

    1(lowest level) 2 3 4 5 6 7 (highest level)
12. İngilizce öğrenmeye başladığınız yaş: *(Onset age of English education)*
13. Genel olarak dinlediğiniz müzik türünü, Doğu - Batı müziği olarak değerlendirecek olursanız nereye koyardınız? *(Where would your musical*

preference reside on a scale at Eastern music at one end and Western music on the other?)

1(Doğu müziği) 2 3 4 5 6 7 (Batı müziği)

1(Eastern music) 2 3 4 5 6 7 (Western music)

14. Bir müzik enstrümanı çalıyor veya şarkı söylüyor musunuz?

- Evet *(Yes)*
- Hayır *(No)*

15. Müziğe ilginiz var mı?

- Evet *(Yes)*
- Hayır *(No)*

Goldsmiths Musical Sophistication Index (Gold-MSI)

Kesinlikle katılmıyorum 1 2 3 4 5 6 7 Kesinlikle katılıyorum

(Completely disagree) 1 2 3 4 5 6 7 (Completely agree)

1. Boş zamanlarımın çoğunu müzikle alakalı aktiviteler yaparak geçiririm. *(I spend a lot of my free time doing music-related activities.)*
2. Bazen tüylerimi diken diken eden müzikler dinlerim. *(I sometimes choose music that can trigger shivers down my spine.)*
3. Müzik hakkında internette yazılar yazmaktan hoşlanırım. *(I enjoy writing about music, for example on blogs and forums.)*
4. Birisi bir şarkı söylediğinde şarkıyı bilmesem bile eşlik edebilirim. *(If somebody starts singing a song I don't know, I can usually join in.)*

5. Birinin iyi bir şarkıcı olup olmadığını anlayabilirim. *(I am able to judge whether someone is a good singer or not.)*
6. Bir şarkıyı ilk defa dinlediğimi genellikle fark edebilirim. *(I usually know when I'm hearing a song for the first time.)*
7. Hafızamdan şarkı söyleyebilir veya müzik yapabilirim. *(I can sing or play music from memory.)*
8. Aşina olmadığım müzik türleri ilgimi çeker ve onları daha fazla öğrenmek isterim. *(I'm intrigued by musical styles I'm not familiar with and want to find out more.)*
9. Müzik parçaları bende nadiren duygu uyandırır. *(Pieces of music rarely evoke emotions for me.)*
10. Kayıtta şarkı söylerken doğru sesleri çıkarabilirim. *(I am able to hit the right notes when I sing along with a recording.)*
11. Bir şarkının tonunu bilsem bile hataları fark etmekte zorlanırım. *(I find it difficult to spot mistakes in a performance of a song even if I know the tune.)*
12. Aynı şarkının farklı yorumları arasındaki farkları anlayabilir ve tartışabilirim. *(I can compare and discuss differences between two performances or versions of the same piece of music.)*
13. Farklı bir yorumcu tarafından icra edildiğinde bildiğim bir şarkıyı tanımakta zorlanırım. *(I have trouble recognizing a familiar song when played in a different way or by a different performer.)*
14. Müzik performans yeteneklerimle ilgili olarak hiç iltifat almadım. *(I have never been complimented for my talents as a musical performer.)*
15. Sık sık internette müzik ile ilgili şeyler arar ve okurum. *(I often read or search the internet for things related to music.)*

16. Kendimi motive etmek veya heyecanlandırmak için belli bir müzik türü seçerim. *(I often pick certain music to motivate or excite me.)*
17. Birisi bildiğim bir tonda şarkı söylerken ona tonal olarak uyumlu bir şekilde eşlik edemem. *(I am not able to sing in harmony when somebody is singing a familiar tune)*
18. İnsanlar müzik yaparken tonun veya ritmin dışına çıktıklarında bunu fark edebilirim. *(I can tell when people sing or play out of time with the beat.)*
19. Bir müzik parçasında neyin özel olduğunu tanımlayabilirim. *(I am able to identify what is special about a given musical piece.)*
20. Bir müzik parçasının bende uyandırdığı duygular hakkında konuşabilirim. *(I am able to talk about the emotions that a piece of music evokes for me.)*
21. Paramı müziğe harcamam. *(I don't spend much of my disposable income on music.)*
22. İnsanların müzik yaparken tonun dışına çıktıklarını (detone olduklarını) anlayabilirim. *(I can tell when people sing or play out of tune.)*
23. Şarkı söylerken tonun dışına çıkıp çıkmadığım hakkında hiçbir fikrim olmaz. *(When I sing, I have no idea whether I'm in tune or not.)*
24. Müzik benim için bir tür bağımlılıktır - onsuz yaşayamam. *(Music is kind of an addiction for me - I couldn't live without it.)*
25. Topluluk içinde şarkı söylemeyi sevmem çünkü yanlış notaları çıkaracağımdan korkarım. *(I don’t like singing in public because I’m afraid that I would sing wrong notes.)*
26. Bir müzik parçası dinlediğimde genellikle türünü (genre) ayırt edebilirim. *(When I hear a music I can usually identify its genre.)*

27. Kendimi bir müzisyen olarak tanımlamam. *(I would not consider myself a musician.)*
28. Karşıma yeni çıkan müzikleri takip ederim (örn: yeni çıkan sanatçı veya albümler). *(I keep track of new of music that I come across (e.g. new artists or recordings))*
29. Bir şarkıyı iki-üç defa dinledikten sonra genellikle söyleyebilirim. *(After hearing a new song two or three times, I can usually sing it by myself.)*
30. Bir şarkıyı ilk defa duymam, birkaç saat sonra söylemem için yeterlidir. *(I only need to hear a new tune once and I can sing it back hours later.)*
31. Müzik geçmişteki insan ve mekanlara dair anılarımı tetikleyebilir. *(Music can evoke my memories of past people and places.)*
32. ___ yıldır her gün vokal veya enstrüman pratiği yapıyorum. *(I engaged in regular, daily practice of a musical instrument (including voice) for ___ years.)*
33. İlgimin en yüksek olduğu zamanlarda, günde ___ saat birincil enstrümanımla pratik yaptım. *(At the peak of my interest, I practiced ___ hours per day on my primary instrument.)*
34. Geçtiğimiz 12 ayda, ___ canlı müzik etkinliğinde dinleyici olarak bulundum. *(I have attended _ live music events as an audience member in the past twelve months.)*
35. ___ yıl formal müzik eğitimi aldım. *(I have had formal training in music theory for __ years.)*
36. Hayatım boyunca ___ yıl enstrüman veya vokal eğitimi aldım. *(I have had __ years of formal training on a musical instrument (including voice) during my lifetime.)*

37. ___ müzik enstrümanı çalabiliyorum. *(I can play ___ musical instruments)*

38. Günde ___ müzik dinliyorum. *(I listen attentively to music for __ per day.)*

- En iyi çaldığım enstrüman: (eğer enstrüman çalmıyor ama şarkı söylüyorsanız "vokal" olarak belirtiniz) *(The instrument I play most proficiently (if you don't play an instrument but sing, please specify it as "vocals"))*
- Katıldığınız bu araştırma hakkında bir yorumunuz, istek veya öneriniz varsa lütfen aşağıya yazarak bizimle paylaşın. *(Please share your comments and requests below)*
- Araştırmanın sonuçlarına dair bilgi almak istiyorsanız e-mail adresinizi yazınız. *(Please enter your e-mail address if you would like to receive information regarding the results of the study)*

APPENDIX B

OTHER RESULTS

Table B1. Average Proportions of Iambic Responses for all Confidence Ratings in Experiment 1A

				All responses		Responses rated as "somewhat sure" and very sure"		Responses rated as "very sure"	
Manipulation	Ratio	Base tone duration	Order of presentation	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
Duration				0.68	0.02	0.69	0.02	0.75	0.03
Amplitude				0.51	0.03	0.48	0.03	0.45	0.05
Pitch				0.53	0.02	0.54	0.03	0.48	0.04
Duration	1.25	150 ms	forward	0.60	0.06	0.57	0.07	0.62	0.11
Duration	1.25	150 ms	reversed	0.57	0.06	0.61	0.07	0.58	0.10
Duration	1.25	250 ms	forward	0.51	0.06	0.49	0.07	0.42	0.09
Duration	1.25	250 ms	reversed	0.54	0.06	0.55	0.06	0.51	0.09
Duration	1.75	150 ms	forward	0.72	0.06	0.75	0.06	0.76	0.07
Duration	1.75	150 ms	reversed	0.62	0.06	0.65	0.07	0.73	0.08
Duration	1.75	250 ms	forward	0.80	0.05	0.82	0.05	0.82	0.06
Duration	1.75	250 ms	reversed	0.74	0.06	0.76	0.06	0.78	0.07
Duration	3	150 ms	forward	0.85	0.05	0.85	0.05	0.91	0.04
Duration	3	150 ms	reversed	0.77	0.05	0.79	0.05	0.83	0.06
Duration	3	250 ms	forward	0.75	0.05	0.77	0.05	0.85	0.05
Duration	3	250 ms	reversed	0.69	0.06	0.68	0.06	0.74	0.06
Amplitude	1.5	150 ms	forward	0.57	0.06	0.60	0.07	0.71	0.11
Amplitude	1.5	150 ms	reversed	0.48	0.06	0.49	0.07	0.42	0.10
Amplitude	1.5	250 ms	forward	0.48	0.06	0.49	0.07	0.43	0.11
Amplitude	1.5	250 ms	reversed	0.45	0.06	0.41	0.06	0.38	0.09
Amplitude	2	150 ms	forward	0.51	0.06	0.51	0.07	0.45	0.11
Amplitude	2	150 ms	reversed	0.54	0.06	0.48	0.07	0.53	0.13
Amplitude	2	250 ms	forward	0.48	0.06	0.48	0.07	0.41	0.09
Amplitude	2	250 ms	reversed	0.55	0.06	0.51	0.07	0.38	0.10
Pitch	1.25	150 ms	forward	0.60	0.06	0.60	0.07	0.52	0.10
Pitch	1.25	150 ms	reversed	0.54	0.06	0.54	0.06	0.56	0.10
Pitch	1.25	250 ms	forward	0.65	0.06	0.69	0.06	0.67	0.10
Pitch	1.25	250 ms	reversed	0.52	0.06	0.57	0.07	0.55	0.09
Pitch	1.5	150 ms	forward	0.40	0.06	0.43	0.07	0.27	0.10
Pitch	1.5	150 ms	reversed	0.49	0.06	0.50	0.07	0.57	0.11
Pitch	1.5	250 ms	forward	0.49	0.06	0.47	0.07	0.39	0.08
Pitch	1.5	250 ms	reversed	0.52	0.06	0.53	0.07	0.43	0.10
Pitch	2	150 ms	forward	0.63	0.06	0.65	0.07	0.50	0.11
Pitch	2	150 ms	reversed	0.51	0.06	0.54	0.07	0.43	0.11
Pitch	2	250 ms	forward	0.48	0.06	0.47	0.07	0.44	0.09
Pitch	2	250 ms	reversed	0.55	0.06	0.53	0.07	0.55	0.09

Table B2. Average Proportions of Iambic Responses for all Confidence Ratings in Experiment 1B

Manipulation	Ratio	Base tone duration	Order of presentation	All responses		Responses rated as "somewhat sure" and very sure"		Responses rated as "very sure"	
				Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
Duration				0.66	0.03	0.68	0.03	0.70	0.04
Amplitude				0.47	0.03	0.46	0.04	0.45	0.05
Pitch				0.55	0.03	0.55	0.04	0.54	0.04
Duration	1.25	150 ms	forward	0.60	0.04	0.62	0.05	0.63	0.07
Duration	1.25	150 ms	reversed	0.51	0.05	0.50	0.05	0.53	0.07
Duration	1.25	250 ms	forward	0.60	0.04	0.59	0.04	0.59	0.05
Duration	1.25	250 ms	reversed	0.56	0.04	0.57	0.05	0.55	0.06
Duration	1.75	150 ms	forward	0.66	0.04	0.68	0.04	0.62	0.06
Duration	1.75	150 ms	reversed	0.67	0.04	0.67	0.04	0.68	0.06
Duration	1.75	250 ms	forward	0.74	0.04	0.74	0.04	0.75	0.04
Duration	1.75	250 ms	reversed	0.66	0.05	0.66	0.05	0.69	0.06
Duration	3	150 ms	forward	0.78	0.04	0.78	0.04	0.78	0.05
Duration	3	150 ms	reversed	0.74	0.04	0.76	0.04	0.76	0.05
Duration	3	250 ms	forward	0.78	0.04	0.78	0.04	0.80	0.05
Duration	3	250 ms	reversed	0.69	0.04	0.70	0.04	0.69	0.05
Amplitude	1.5	150 ms	forward	0.51	0.04	0.48	0.05	0.46	0.07
Amplitude	1.5	150 ms	reversed	0.46	0.05	0.46	0.05	0.44	0.07
Amplitude	1.5	250 ms	forward	0.47	0.05	0.44	0.05	0.42	0.06
Amplitude	1.5	250 ms	reversed	0.45	0.05	0.45	0.05	0.41	0.06
Amplitude	2	150 ms	forward	0.51	0.04	0.47	0.05	0.47	0.07
Amplitude	2	150 ms	reversed	0.50	0.04	0.49	0.04	0.48	0.06
Amplitude	2	250 ms	forward	0.40	0.04	0.39	0.05	0.38	0.06
Amplitude	2	250 ms	reversed	0.47	0.05	0.47	0.05	0.41	0.06
Pitch	1.25	150 ms	forward	0.58	0.04	0.61	0.05	0.66	0.06
Pitch	1.25	150 ms	reversed	0.51	0.04	0.49	0.05	0.47	0.06
Pitch	1.25	250 ms	forward	0.54	0.05	0.49	0.05	0.55	0.06
Pitch	1.25	250 ms	reversed	0.50	0.05	0.51	0.05	0.48	0.06
Pitch	1.5	150 ms	forward	0.57	0.04	0.58	0.05	0.54	0.07
Pitch	1.5	150 ms	reversed	0.51	0.04	0.49	0.05	0.48	0.07
Pitch	1.5	250 ms	forward	0.59	0.04	0.59	0.05	0.58	0.06
Pitch	1.5	250 ms	reversed	0.55	0.05	0.54	0.05	0.50	0.07
Pitch	2	150 ms	forward	0.58	0.04	0.57	0.05	0.60	0.06
Pitch	2	150 ms	reversed	0.52	0.05	0.49	0.06	0.46	0.07
Pitch	2	250 ms	forward	0.60	0.05	0.58	0.05	0.54	0.07
Pitch	2	250 ms	reversed	0.54	0.05	0.52	0.05	0.53	0.06

Table B3. Average Proportions of Iambic Responses for all Confidence Ratings in Experiment 2.

				All responses		Responses rated as "somewhat sure" and very sure"		Responses rated as "very sure"	
Manipulation	Ratio	Syllable	Order of presentation	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
Duration				0.73	0.03	0.75	0.03	0.80	0.03
Amplitude				0.37	0.03	0.35	0.04	0.30	0.05
Pitch				0.57	0.04	0.57	0.04	0.59	0.06
Duration	1.25	[gü]	forward	0.57	0.06	0.52	0.06	0.60	0.08
Duration	1.25	[gü]	reversed	0.55	0.05	0.55	0.05	0.55	0.08
Duration	1.25	[kı]	forward	0.54	0.05	0.55	0.06	0.62	0.08
Duration	1.25	[kı]	reversed	0.56	0.05	0.54	0.05	0.59	0.08
Duration	1.75	[gü]	forward	0.81	0.04	0.83	0.04	0.85	0.05
Duration	1.75	[gü]	reversed	0.77	0.05	0.79	0.05	0.84	0.06
Duration	1.75	[kı]	forward	0.73	0.05	0.74	0.05	0.79	0.05
Duration	1.75	[kı]	reversed	0.78	0.05	0.79	0.05	0.80	0.06
Duration	3	[gü]	forward	0.93	0.02	0.93	0.03	0.93	0.04
Duration	3	[gü]	reversed	0.86	0.04	0.86	0.04	0.86	0.05
Duration	3	[kı]	forward	0.84	0.05	0.85	0.04	0.89	0.05
Duration	3	[kı]	reversed	0.83	0.04	0.84	0.04	0.85	0.05
Amplitude	1.5	[gü]	forward	0.38	0.05	0.38	0.05	0.31	0.08
Amplitude	1.5	[gü]	reversed	0.40	0.05	0.40	0.05	0.38	0.08
Amplitude	1.5	[kı]	forward	0.44	0.05	0.39	0.06	0.40	0.09
Amplitude	1.5	[kı]	reversed	0.34	0.04	0.29	0.05	0.25	0.08
Amplitude	2	[gü]	forward	0.42	0.06	0.42	0.07	0.42	0.09
Amplitude	2	[gü]	reversed	0.33	0.05	0.29	0.05	0.27	0.07
Amplitude	2	[kı]	forward	0.39	0.06	0.40	0.06	0.23	0.07
Amplitude	2	[kı]	reversed	0.26	0.05	0.25	0.05	0.21	0.07
Pitch	1.25	[gü]	forward	0.58	0.06	0.59	0.06	0.65	0.08
Pitch	1.25	[gü]	reversed	0.51	0.05	0.50	0.06	0.49	0.09
Pitch	1.25	[kı]	forward	0.54	0.06	0.53	0.06	0.59	0.09
Pitch	1.25	[kı]	reversed	0.51	0.06	0.48	0.06	0.57	0.09
Pitch	1.5	[gü]	forward	0.63	0.06	0.62	0.06	0.68	0.08
Pitch	1.5	[gü]	reversed	0.60	0.06	0.60	0.07	0.63	0.08
Pitch	1.5	[kı]	forward	0.58	0.05	0.62	0.06	0.66	0.09
Pitch	1.5	[kı]	reversed	0.56	0.05	0.59	0.06	0.62	0.08
Pitch	2	[gü]	forward	0.58	0.05	0.57	0.05	0.64	0.07
Pitch	2	[gü]	reversed	0.63	0.05	0.63	0.05	0.69	0.07
Pitch	2	[kı]	forward	0.54	0.06	0.54	0.07	0.60	0.09
Pitch	2	[kı]	reversed	0.53	0.05	0.52	0.06	0.60	0.09

REFERENCES

Altmann, H. (2006). The perception and production of second language stress: A cross-linguistic experimental study. Newark, DE: University of Delaware.

Bhatara, A., Boll-Avetisyan, N., Agus, T., Höhle, B., & Nazzi, T. (2016). Language experience affects grouping of musical instrument sounds. *Cognitive Science*, *40*(7), 1816-1830. doi:10.1111/cogs.12300

Bhatara, A., Boll-Avetisyan, N., Unger, A., Nazzi, T., & Höhle, B. (2013). Native language affects rhythmic grouping of speech. *Journal of the Acoustical Society of America*, *134*(5), 3828-3843. doi:10.1121/1.4823848

Bion, R. A. H., Benavides-Varela, S., & Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Language and Speech*, *54*(1), 123-140. doi:10.1177/0023830910388018

Bolton, T. L. (1894). Rhythm. *The American Journal of Psychology*, *6(*2), 145-238. doi:10.2307/1410948

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, *14*(4), 362-366. doi:10.1111/1467-9280.24441

Charette, M. (2008). The vital role of the trochaic foot in explaining Turkish word endings. *Lingua*, *118*(1), 46-65.

Crowhurst, M. J. (2018). The influence of varying vowel phonation and duration on rhythmic grouping biases among Spanish and English speakers. *Journal of Phonetics*, *66*, 82-99. doi:10.1016/j.wocn.2017.09.001

Crowhurst, M. J., & Teodocio Olivares, A. (2014). Beyond the iambic-trochaic law: The joint influence of duration and intensity on the perception of rhythmic speech. *Phonology*, *31*(1), 51-94. doi:10.1017/S0952675714000037

De la Mora, Daniela M, Nespor, M., & Toro, J. M. (2013). Do humans and nonhuman animals share the grouping principles of the iambic–trochaic law? *Attention, Perception, & Psychophysics, 75*(1), 92-100. doi:10.3758/s13414-012-0371-3

Deutsch, D. (1991). The tritone paradox: An influence of language on music perception. *Music Perception: An Interdisciplinary Journal*, *8*(4), 335-347.

Dupoux, E., Pallier, C., Sebastián, N., & Mehler, J. (1997). A destressing “deafness” in French? *Journal of Memory and Language*, *36*, 406–421.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress “deafness”. *The Journal of the Acoustical Society of America*, *110*(3), 1606-1618.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress ‘deafness’: The case of French learners of Spanish. *Cognition*, *106*(2), 682-706.

Göksel, A., & Kerslake, C. (2004). Turkish: A comprehensive grammar. *Routledge*.

Hay, J. S. F., & Diehl, R. L. (2007). Perception of rhythmic grouping: Testing the iambic/trochaic law. *Perception and Psychophysics*, *69*(1), 113-122. doi:10.3758/BF03194458

Hay, J. F., & Saffran, J. R. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy*, *17*(6), 610-641. doi:10.1111/j.1532-7078.2011.00110.x

Hayes, B. (1995). *Metrical stress theory: Principles and case studies.* Chicago: University of Chicago Press.

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior and Development*, *32*(3), 262-274. doi:10.1016/j.infbeh.2009.03.004

Hunyadi, L., (2015). Grouping, symmetry, and rhythm in language. In Ralf Vogel, Ruben van de Vijver (Eds), *Rhythm in cognition and grammar: A Germanic perspective.* (pp. 17-53). Berlin, Germany: Walter de Gruyter GmbH.

Hyde, B. (2011). The Iambic—Trochaic Law In M. Oostendorp, C. J. Ewen, E. Hume and K. Rice (Eds.), *The Blackwell Companion to Phonology* (pp. 1-26). Chichester: Wiley-Blackwell

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America*, *124*(4), 2263-2271. doi:10.1121/1.2973189

Jacoby, N., & McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*, *27*(3), 359-370.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*(1), 1–23.

Kabak, B. (2016). Refin(d)ing Turkish stress as a multifaceted phenomenon. Presented at the Second Conference on Central Asian Languages and Linguistics (ConCALL-2), Indiana University

Kabak, B., & Vogel, I. (2001). The phonological word and stress assignment in Turkish. *Phonology*, *18*(3), 315-360.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception. 36*(ECVP Abstract Supplement):14. Whats New, 36, 14.

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences*, *22*(10), 896-910. doi:10.1016/j.tics.2018.08.002

Langus, A., Mehler, J., & Nespor, M. (2017). Rhythm in language acquisition. *Neuroscience and Biobehavioral Reviews*, *81*(Pt B), 158-166. doi:10.1016/j.neubiorev.2016.12.012

Langus, A., Seyed-Allaei, S., Uysal, E., Pirmoradian, S., Marino, C., Asaadi, S., Nespor, M. (2016). Listening natively across perceptual domains? *Journal of Experimental Psychology: Learning Memory and Cognition*, 42(7), 1127-1139. doi:10.1037/xlm0000226

Lefebvre, H. (2004). *Rhythmanalysis: Space, time, and everyday life*. London: Continuum.

Levi, S. V. (2005). Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association, 35*(1), 73-97. doi:10.1017/S0025100305001921

Lu, S., Vigário, M., Correia, S., Jerónimo, R., & Frota, S. (2018). Revisiting stress "deafness" in European Portuguese–a behavioral and ERP study. *Frontiers in Psychology*, *9*, 2486.

Molnar, M., Carreiras, M., & Gervain, J. (2016). Language dominance shapes non-linguistic rhythmic grouping in bilinguals. *Cognition*(*152*), 150-159.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart L. (2014). The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLoS ONE, 9*(2): e89642. doi:10.1371/journal.pone.0089642

Nespor, M., Vijver, R. v. d., Schraudolf, H., Shukla, M., Avesani, C., & Donati, C. (2008). Different phrasal prominence realizations in VO and OV languages. *Lingue e Linguaggio, 7(*2), 139-167.

Nespor, M., Shukla, M., & Mehler, J. (2011). Stress-timed vs. syllable-timed languages. In M. Van. Oostendorp, C. J. Ewen, E. Hume & K. Rice (eds.), *The Blackwell Companion to Phonology* (pp. 1147-1159). Oxford: Wiley-Blackwell.

Özçelik, Ö. (2017). The Foot is not an obligatory constituent of the Prosodic Hierarchy: "Stress" in Turkish, French and child English. *The Linguistic Review, 34*(1), 157-213.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674.

Peperkamp, S., & Dupoux, E. (2002). A typological study of stress "deafness.". *Laboratory Phonology*, *7*, 203-40.

Polak, R., Jacoby, N., Fischinger, T., Goldberg, D., Holzapfel, A., & London, J. (2018). Rhythmic prototypes across cultures: a comparative study of tapping synchronization. *Music Perception: An Interdisciplinary Journal*, *36*(1), 1-23.

Pycha, A. (2006). A duration-based solution to the problem of stress realization in Turkish. *UC Berkeley PhonLab Annual Report, 2*(2).

Rahmani, H., Rietveld, T., & Gussenhoven, C. (2015). Stress “deafness” reveals absence of lexical marking of stress or tone in the adult grammar. *PloS one*, *10*(12), e0143968.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265-292. doi:10.1016/S0010-0277(99)00058-X

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928. doi:10.1126/science.274.5294.1926

Schmidt-Kassow, M., Rothermich, K., Schwartze, M., & Kotz, S. A. (2011). Did you get the beat? Late proficient French-German learners extract strong–weak patterns in tonal but not in linguistic sequences. *Neuroimage*, *54*(1), 568-576. doi:10.1016/j.neuroimage.2010.07.062

Sezer, E. (1981). On non-final stress in Turkish. *Journal of Turkish studies*, *5*(61-69).

Snape, N., del Pilar García Mayo, M., and Gürel, A. (2009). Spanish, Turkish, Japanese and Chinese L2 learners’ acquisition of generic reference. In Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009), ed. Melissa Bowles et al., 1-8. Somerville, MA: Cascadilla Proceedings Project.

Toro, J. M., & Nespor, M. (2015). Experience-dependent emergence of a grouping bias. *Biology Letters*, 11(9), 20150374.

Trainor, L. J., & Adams, B. (2000). Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception and Psychophysics*, *62*(2), 333-340. doi:10.3758/BF03205553

Yeung, H., Bhatara, A., & Nazzi, T. (2018). Learning a phonological contrast modulates the auditory grouping of rhythm. *Cognitive Science*, *42*(6), 2000-2020. doi:10.1111/cogs.12635

Yoshida, K. A., Iversen, J. R., Patel, A. D., Mazuka, R., Nito, H., Gervain, J., & Werker, J. F. (2010). The development of perceptual grouping biases in infancy: A Japanese-English cross-linguistic study. *Cognition*, *115*(2), 356-361. doi:10.1016/j.cognition.2010.01.005

Zora, H., Heldner, M., Schwarz, I. (2016). Perceptual correlates of Turkish word stress and their contribution to automatic lexical access: Evidence from early ERP components. *Frontiers in Neuroscience*, *10*, 7. doi:10.3389/fnins.2016.00007