ACOUSTIC CORRELATES OF LEXICAL STRESS IN NATIVE SPEAKERS OF UYGHUR AND L2 LEARNERS

BY

MAHIRE YAKUP (YAKEFU MAYILA)

Submitted to the graduate degree program in Linguistics

and the Graduate Faculty of the University of Kansas in partial fulfillment of

the requirements for the degree of Doctor of Philosophy.

____

Co-chairperson Dr. Joan Sereno

____

Co-chairperson Dr. Arienne M Dwyer

____

Dr. Allard Jongman

____

Dr. Jie Zhang

____

Dr. Carlos M Nash

Date Defended: 6 March 2013

The Dissertation Committee for MAHIRE YAKUP (YAKEFU MAYILA)

certifies that this is the approved version of the following dissertation:

ACOUSTIC CORRELATES OF LEXICAL STRESS IN NATIVE SPEAKERS OF UYGHUR AND L2 LEARNERS

Co-Chairperson Dr. Joan Sereno

Co-Chairperson Dr. Arienne M Dwyer

Date Approved: 6 March 2013

Abstract

Some syllables are louder, longer and stronger than other syllables at the lexical level. These prominent prosodic characteristics of certain syllables are captured by suprasegmental features including fundamental frequency, duration and intensity. A language like English uses fundamental frequency, duration and intensity to distinguish stressed syllables from unstressed syllables; however, a language like Japanese only uses fundamental frequency to distinguish the stressed syllables from unstressed syllables.

This study investigates the stress pattern of Uyghur, a Turkic language, as produced by native and non-native speakers. The first three experiments provide a detailed phonetic analysis in order to determine the acoustic cues to stress in Uyghur. In Experiment 1, six disyllabic minimal pairs (e.g., *A-cha*, *a-CHA*), contrasting in location of stress, were produced by five native Uyghur speakers with three repetitions in a fixed sentence context. In order to generalize the results from the small set of minimal pairs in the first experiment, Experiment 2 examined the initial syllable of disyllabic nouns that contrasted in first-syllable stress (e.g., *DA-ka*, *da-LA*) while syllabic structure (CV versus CVC) was also manipulated. In both experiments, average fundamental frequency, syllable duration, and average intensity were collected in accented and unaccented syllables. The results from both experiments showed that there were significant differences in duration and intensity between stressed and unstressed syllables, with the intensity differences moderated by syllable structure. No difference was found in fundamental frequency. Experiment 3 investigated the role of F0 in lexical stress. Experiment 3 focused on the interaction between sentential intonation and lexical stress in which the declarative assertion sentence (falling F0) and the declarative question sentence (rising F0) were used. The results confirmed the previous experiments. No interaction between sentential intonation and lexical

stress indicated that the obtained duration effect was due to lexical stress. There were no effects of fundamental frequency or intensity in terms of stress. While previous studies have classified Uyghur as a pitch-accent and a stress-accent language, the present acoustic data suggest that native speakers make no use of pitch cues to signal stress in Uyghur.

Previous research has focused on the acquisition of lexical stress by non-native speakers of English. This study also examined the acquisition of lexical stress by English learners of Uyghur. Five highly advanced English learners of Uyghur produced the six minimal pairs and disyllabic nouns contrasting in the first syllables. The stimuli that were produced by L2 learners were the same as in Experiment 1 and Experiment 2. Highly advanced Uyghur learners used duration as a cue and did not use fundamental frequency and intensity as stress cues. The results indicated that native-like lexical stress can be acquired at the high advanced level.

Acknowledgments

I have been waiting for this day for a long time. Now that I finished my dissertation and presentation, I would like to thank my chairpersons and committee members. Since the beginning of this work, I have been keeping all these people in my mind. I appreciate their generous help and support. Without their help, instruction, revisions, and comments, I would not be here. I am very blessed to have these professors as my committee members. Everything is so clear and so fresh in my mind now and will be throughout the rest of my life. I enjoyed a good experience as a graduate student with them.

First, I would like to send my deep thanks to my advisor Dr. Joan Sereno for her support, encouragement, and mentorship. This dissertation would not have been possible without her generous support and step-by step guidance. She was my committee member for my master's thesis and co-chair for my doctoral dissertation. When I made my decision about the turning point in my graduate career, she stood by my side and gave me a hand that made this journey successful at the end. The first time when I discussed what I should do in my dissertation, she listed all the possibilities and discussed the pros and cons with me in detail. When I picked the topic, she was delighted and helped me through each step. She put every piece of information in the right position and provided me with detailed feedback. Her office was always open to me and we had many comfortable and progressive discussions on my minor and major projects and dissertation. I appreciated her fast responses and her readiness to take time to discuss my thesis with me. I have benefited greatly from her knowledge and experience in research. In her class I got great opportunities to practice my thesis topic and organize my ideas. Every time, I was encouraged and fulfilled with new energy from working with her.

At the same time, I would like to thank my co-chair Dr. Arienne Dwyer. She opened the door for me to study linguistics at the University of Kansas. I changed from a pure psychologist into a linguist at this university. Without her help and encouragement, I would not be here today; I learned a lot from her. She is leading me to become as Turkologist and helping me to preserve my language (Uyghur) through training me to do research as one from my own community. When we were collaborating on the Uyghur textbook, I learned formal linguistics outside of the Linguistic Department. I had opportunities as a language consultant at Infield 2010 and as a teaching assistant at CoLang 2012, which provided the chance to learn language revitalization and documentation practices. She is not only excited by linguistic phenomena, but is eager to delve into the work. These characteristics influenced me to become a Turkologist and focus on my language as my dissertation topic. She is like my 'elder sister' in my personal life but she is an encouraging mentor in my academic life. She always gave me good advice and helped me smoothly adjust to both aspects of my life in America. Both co-chairs were wonderful to work with. Their advice will lead me toward being a good scholar and researcher in the future.

My deep gratitude also goes to Dr. Allard Jongman. His frequent assistance and guidance were invaluable to me. He always surprises me with his deep knowledge of acoustic phonetics. His fast, to-the-point, and detailed responses are always amazing and much appreciated. Whenever I was confused about the methods or logic of the research, he stood up and gave me an explanation. When I was excited to do more about related issues, he put me on the right track and explained to me how much I can do and how far I can go. With his advice, I focused on my research in detail and stayed within the scope of my topic.

I also would like to send my thanks to Dr. Zhang Jie, who is my committee member, for his thoughtful input concerning my dissertation. I appreciated his guidance and help with my

introduction. Among the graduate students, he is well known for his fast grading and providing feedback on final or term papers. His critical thinking always inspired me. Also, many thanks go to Dr. Carlos Nash. His encouragement and jokes always relieve me. Even though I did not take a course from him, I knew him when I was at Infield 2010 and became better acquainted with him after he came to the University of Kansas.

I had lots of benefits from Linguistics 850 and classmates including Kelly Berkson, Steve Poltzer-Ahles, Liu Jiang, Hyunjung Lee, and Goon Lee. Whenever I presented my data and ideas, these colleagues gave me suggestions and comments. All these comments and suggestions helped me move forward. In addition, I am also thankful for Dr. Christina Jensen for her assistance in writing my dissertation. She helped me to use correct English from and expressions. At the last minute I wanted to add the etymology of my data, at which time Turki BinTurki and Hesham Aldamen checked the Arabic source of Uyghur words. I appreciate their kindness and help. I also thank current graduate students such as Jon Coffee and Travis Major for discussing the Uyghur language. Working with them makes me happy and inspired to be a linguist. I also appreciated Abduwali Ayup and Gulnar Eziz for being pre-test participants and commenting on my experiments. I also thank Dr. Meghripet Kenji who is a professor at Xinjiang Normal University, who helped collect data from Xinjiang. I also appreciate all my participants. Without their participation in this research, I could not have completed my project. I would like to extend my thanks to all the people who have helped me individually.

I also express my deep gratitude to the EALC department and Center for East Asian Studies for financially supporting me as a graduate teaching assistant for these years. Without their support I could not stay a single day in this graduate program. I also appreciate the generous help of the Graduate School for Summer Research Fund for allowing me to focus on writing and collecting data without extra work.

A final special thank you note goes to my mother, Rizwan Ali. She always stands by my side to support me. Without her encouragement and insistence, I would not be here today.

Dedication

To my parents

Rizwan Ali

Yakup Höseyin (1937-1992)

Contents

Tables

Figures

Chapter 1 Introduction

1.1 Theory background of the categorization of languages

Acoustic parameters provide the information that certain syllables are more prominent than others within a word. The prominent prosodic characteristics of certain syllables are captured by suprasegmental features such as fundamental frequency, duration, intensity and vowel quality (Jun, 2005).

Prominence of syllables is a relative term where prominence exists in only comparing the environment in which the “prominent” syllables occur. The environment could be within words or across words. For example, in English, people perceive “SUB” in “SUBject” (upper case is the stressed syllable) as higher, louder, and longer compared to the “ject” in “SUBject”; thus, English native speakers perceive “SUBject” as a noun. If they perceive “JECT” in “subJECT” as higher, louder and longer than the “sub” in “subJECT”, then they perceive it as a verb. Usually, researchers use the minimal pairs such as “SUBject” versus “subJECT” to examine the acoustic parameters of stress. Acoustically, “SUB” in “SUBject” has a higher fundamental frequency, a longer duration, and a stronger intensity compared to the parameters of “sub” in “subJECT.” It is also the case that the “JECT” in “subJECT” has higher fundamental frequency, longer duration and stronger intensity compared to the “ject” in “SUBject.” Therefore, there is a match between human perception and these acoustic realizations of prominence for this English example.

The prominence of syllables is language-specific. Languages may use all, or some of the acoustic parameters, to indicate their prominent syllables. Speakers perceive the prominent syllables based on pitch, duration, or intensity of syllables. For example, some languages, such as English, use the combination of fundamental frequency, duration, and intensity of syllables for assigning the prominent syllables. On the other hand, other languages only choose a subset of these acoustic parameters (such as fundamental frequency) to assign prominent syllables. Therefore, languages can be classified based on acoustic parameters that are used by speakers to distinguish prominent syllables from non-prominent syllables.

Languages can be categorized as stress-accent languages, tone-accent languages, or pitch-accent languages, based on the characteristics of the prominent syllables (Beckman, 1986; Jun 2005). Though some researchers (Halle & Idsardi, 1994; Hayes, 1995) use phonological criteria, others (Beckman, 1986; Levi, 2005; Sluijter & Van Heuven, 1996a, b; Van Der Mark, 2003) use phonetic cues to distinguish the stress types of languages. Beckman (1986) and Hualde et al. (2002) support this three-way categorization, in which they define pitch-accent languages as a category independent from stress-accent and tone-accent languages. Hyman (2006) suggests that there may be only two prototypes of languages: stress-accent versus tone-accent languages while pitch-accent languages are the combination of tone-accent and stress-accent languages.

The three-way categorization of languages includes criteria based on properties from phonological and phonetic aspects. As shown in Table 1 below, Lindstrom and Remijsen (2005) provided a clear picture of the categorization of languages. It includes both phonological and phonetic aspects of word prosodic typology. The phonological aspects indicate whether a word has contrasted prosodic properties within the word level (syntagmatic) or across the word level (paradigmatic). The phonetic properties of languages focus on the acoustic cues of fundamental frequency, duration, intensity and vowel reduction in prominent syllables. The phonetic aspects include which acoustic parameters distinguish the prominent syllable from non-prominent syllables. In other words, the phonological and phonetic aspects provide us a tool for categorizing the languages in terms of their prosodic properties.

Table 1 The Dimensions of Categorization of Languages

phonology	culminative, syntagmatic characteristics, distinctive function		paradigmatic characteristics
	lexical stress accent	pitch accent	tone accent
	e.g. English	e.g. Japanese	e.g. Mandarin Chinese
minimal pairs	penult SUBject final subJECT	penult KAme 'turtle' final kaME 'jug'	T55 ma 'mother' T35 ma 'hemp' T214 ma 'horse' T51 ma 'scold'
phonetic	use the duration, intensity and vowel quality and other than F0	only use fundamental frequency	

(Lindstrom and Remijsen, 2005; Japanese and English examples adapted from Beckman, 1986)

From Table 1, stress-accent languages and pitch-accent languages share the phonological aspects that both types of languages have a word-prosodic feature that is contrastive at the syntagmatic level. However, these two types of languages differ in phonetic aspects in which stress-accent languages use multiple dimensions, and pitch-accent languages use only pitch as a cue for distinguishing prominent from non-prominent syllables.

According to Ladd (1996), English and Japanese use pitch as an acoustic cue; however, the role of pitch in the two languages differs. In English, pitch is a post-lexical feature, which means that pitch is not lexically specified; on the other hand in Japanese, pitch is lexically specified. Therefore, the origin of pitch is not the same for both English and Japanese. Ladd's model of the word-prosodic typology suggested that the features of pitch are lexical, post-lexical or intonational, as shown in Table 2.

Table 2 Ladd's Categorization of Stress Accent and Non-stress Accent Languages

		Phonetic typology	
		stress accent	non-stress accent
Lexical	Lexical pitch	example: Swedish	example: Japanese
Typology	Post lexical pitch	example: English	example: Bengali

The role of pitch in English, which is at the post-lexical level, is different from that of Japanese, which is at the lexical level as Ladd (Ladd, 1996, p. 156) proposed as below:

Stress and non-stress accent is a phonetic typological dimension (is lexical accent manifested by stress or not?). Lexical versus post lexical specification of pitch features is a phonological phenomenon or even morphological typological dimension (can pitch features be specified in the lexicon or not?)…. If the two typological dimensions are really independent, it ought to be possible to find four types of languages.

Ladd's categorization distinguishes the lexical level F0 from the intonational level F0. In addition, Ladd's model has two advantages. First, his model differentiates stress-accent languages in terms of the function of pitch. In stress-accent languages, pitch can differentiate languages such as English, from languages such as Swedish, both of which are considered stress-accent languages. By distinguishing Accent 1 from Accent 2 in Swedish, Fant and Kruckenberg (1994) found that Accent 1 (HL*H), [high low star, indicates that the first H is words accent, and the L indicates the stressed syllables with star, and then the last H indicates sentence accent rise] was lexically specified in which the stressed syllable has low F0 and the preceding syllable has high tone, and Accent 2 (H*LH) abided by rules, in which the stressed syllables have high tone. In both accent patterns, the sentence contour with falling terminal junctures. This information indicated that pitch in Swedish is a lexical feature rather than an intonational feature as in English. Therefore, Swedish shares the characteristics with Japanese in terms of the origin of pitch, which comes from the lexical level. Secondly, Ladd's model provides phonetic typology and lexical typology of word-prosody. Lexical typology includes the lexical pitch and post-lexical pitch. Ladd's model provides more detailed information about the origin of pitch, which was not clearly discussed in previous categorizations.

In Ladd's model, pitch is one of the indicators in stress assignment. However, his model did not take into account languages such as French, Wolof or Kuot. It is possible that F0 does not play any role in differentiating stress location in some stress-accent languages. For example, in French, pitch is related to syllables in phrases, or pitch is not at the lexical level, and does not have the function of distinguishing stress locations (Dupoux et al. 2008). In Kout, the stress does not include F0 changes. Therefore, Ladd's model cannot account for some languages that do not fit these categories.

Lindstrom and Remijsen (2005), as shown in Table 3 below, modified Ladd's typology of word prosody. According to them, for "no pitch" languages, it is enough to use duration, intensity, vowel quality, and/or a combination of these parameters, instead of using F0, to differentiate the stressed syllables from the unstressed syllables. Therefore, they added another layer onto lexical typology: "no pitch'' languages, in which some stress-accent languages do not use F0, but use the other parameters of duration, intensity, and vowel quality.

On the other hand, non-stress accent languages do not have a no-pitch category, because for non-stress accent/tone languages, pitch is the only cue. By examining the acoustic parameters in the minimal pairs in Kuot, Lindstrom and Remijsen (2005) found that Kuot did not use F0 to distinguish the stressed syllables from the unstressed syllables; instead, Kuot used duration and vowel quality. Therefore, stress-accent languages use other parameters, rather than F0.

Table 3 Lindstrom and Remijsen's (2005) Extension of Ladd's Model

		Phonetic typology stress accent	non-stress accent
Lexical typology	Lexical pitch	example: Swedish	example: Japanese
	Post lexical pitch	example: English	example: Bengali
	No-pitch accent	example: Wolof	example: impossible

Levi (2005) experimentally confirmed that Turkish was a pitch-accent language, rather than a stress-accent language by measuring fundamental frequency, intensity and duration, on 10 near-minimal pairs in noun and verb groups. She found that only pitch had strong correlations to stress locations, compared to other parameters, suggesting that Turkish is similar to Japanese, Basque or Blackfoot in terms of stress. However, she suggests that pitch-accent languages are diverse in terms of the parameters. She did not indicate whether Turkish is close to Bengali or Japanese.

Uyghur is a branch of Turkic languages that includes Turkish. Uyghur may behave like Turkish in which only pitch provides the stress cue, or behave like English using a variety of acoustic parameters. This dissertation will investigate which acoustic parameters provide cues for stress in Uyghur. Knowing the roles of different acoustic parameters is a key to categorizing languages.

***1**.2. Acoustic research on stress-accent languages*

Over half a century, research on stress patterns has focused on acoustic parameters that are strongly correlated to stress assignment. The basic pattern is that people perceive some syllables as stronger, higher or longer than other syllables in speech. Perception of prominent syllables is a psychological process, but these psychological parameters will be realized as F0, duration and intensity, physically. In other words, people use acoustic parameters to see whether syllables are prominent or not in their perception. Therefore, research usually has tested the usefulness of the parameters acoustically, in the perception and production of stress-contrasted words.

1.2.1 Duration, intensity and F0 in stress-accent languages

Stress-accent languages use multiple parameters to assign stress. One of the first to acoustically analyze English stress was Fry (1955). He examined minimal pairs that contrast stress location, but segmentally are similar to each other. For example, the minimal pair SUBject, a noun, and subJECT, a verb, create a minimal pair in English. Twelve English speakers were asked to produce sentences with minimal pairs. He selected the five most effective (participants produce correctly as a noun or a verb) minimal pairs and measured the duration ratios and peak intensity ratios. Based on the plot distribution of duration and intensity in which the x-axis was vowel one and the y-axis was vowel two in each minimal pair, he found that there was very little overlapping of the values for each member in each pair. Even though he used these parameters as test materials for the synthesis of stimuli for a perception of stress task, the results indicated that the stressed syllables had longer duration and stronger intensity than the unstressed syllables in production (Fry, 1955, 1958). In his perception task, he manipulated the duration and intensity ratios by using a step-down or step-up process. 118 participants made judgments about whether the given word was a noun or verb. The participants increased judgments for a noun, when the duration and intensity ratio increased. It was the same for fundamental frequency ratios. Therefore, for English, Fry concluded that duration, intensity and fundamental frequency are cues for stress location.

Further, the usefulness of acoustic parameters was investigated in production experiments. By investigating the same minimal pairs as Fry (1955), Beckman (1986) used the following: ratio analysis (second syllable/first syllable) for fundamental frequency and duration, and for intensity, she used a subtraction (second syllable minus first syllable) analysis. Beckman (1986) confirmed Fry's (1955, 1958) results in which fundamental frequency, duration and intensity are cues for stress assignment in English.

All cues may not work equally well in distinguishing stress location. Some parameters are stronger than others across languages, and some parameters work only for specific languages. Therefore, research has been focusing on the hierarchy of the acoustic parameters that are cues for stress location.

In order to know the hierarchy of the acoustic parameters, a Linear Discriminant Analysis (LDA) provides information about parameters that indicate the most predictable cues for stress locations. LDA usually uses these parameters to predict the stressed and the unstressed syllables, and singles out parameters that correctly classify the stressed syllable from the unstressed syllables. When testing the stress patterns in Dutch and American English, Sluijter and Van Heuven (1996a, 1996b) used LDA for all parameters including duration, pitch (F0), overall intensity, spectral balance or glottal parameters (which indicate the intensity), and vowel quality.

Sluijter and Van Heuven (1996a, 1996b) found that (1) duration is the strongest cue for stress assignment in both languages; (2) spectral balance for Dutch, and glottal parameters for English are placed in the second position in the hierarchy; and (3) vowel quality is in the third position for English, but not for Dutch. However, in both languages, overall intensity and F0 were not strongly correlated to stress. From this research, duration is a strong cue for stress location, compared to intensity and F0. The hierarchy of vowel quality is based on language-specific cues.

The strongest role of duration, in stress assignment, occurs in variety of languages beside English and Dutch. For example, languages such as Spanish and Catalan (Ortega-Llebaria and

Prieto, 2011) use duration as the strongest stress correlate, while pitch and vowel quality were controlled across languages. They confirmed that spectral tilt did not distinguish the stressed syllables from the unstressed syllables, but overall intensity worked for Spanish though not for Catalan.

In addition, polysyllabic Sinhala verbs (Nash, 2005) contrast the stressed syllables from the unstressed syllables by using duration and intensity in the combination in which he assumed that if the syllable durations are equal, only the syllable that had a greater intensity would be perceived as stressed one; on the other hand, if in the equal intensity condition, the syllables that had longer duration would be perceived as stressed one. Therefore, he used a new method in which he multiplied syllable durations with intensity. He could predict that, for two disyllabic words, in the syllables that had equal duration (all CV syllables), and the one that intensity is stronger had stress or vise-versa. Duration was a better predictor than F0 and intensity, in a single parameter. Unlike other research, Nash used a new method, which was the combination of duration and intensity in closed syllables or open syllable words, respectively, as the criterion for stress correlates. He found that 81.7 % of the stressed syllables were correctly classified, by the criterion of duration and intensity, among the 60 stressed syllables marked by a native speaker. Even though the stimuli and languages were different, duration in the stressed syllables is longer than the unstressed syllables.

The hierarchy of intensity and pitch, as a cue in stress assignment also stimulated debates among scholars. Firstly, the debate focuses on the comparison of the role of duration versus intensity in the hierarchy of the acoustic parameters as stress cues. Compared to duration, intensity was not a strong cue. Fry (1955, 1958) indicated that, in the hierarchy order of duration and intensity, duration might be stronger than intensity in the perception study. He controlled duration and manipulated intensity, or controlled intensity and manipulated the duration, in order to know the independent roles of duration and intensity, respectively. He found that the changes in duration had stronger correlates to the stress location and at the same time, the changes in intensity had weaker correlates to the stress in his perception study. Besides English, in Chuvash (Dobrovolsky, 1999), a Turkic language, the duration ratio was a stronger cue than the total amplitude ratio in distinguishing stress location. Therefore, in the hierarchy, duration remains higher than intensity.

Moreover, the debate goes further and focuses on whether intensity is a cue or not for stress assignment. In this situation, the inconsistent results on the role of intensity came from a variety of measurements of intensity, used in different research. The majority of research (Beckman, 1986; Fry, 1955, 1958; Dobrovolsky, 1999) adopted the ratio (second syllable over the first syllable) or total amplitude, average amplitude, and range of amplitude as indicators of intensity.

However, Sluijter and Van Heuven (1996) found that intensity (total amplitude) was not a strong cue for stress location, which was a contradiction to the previous results (Beckman, 1986). Unlike Beckman (1986), who did not tease apart the sentential accent from the stress, Sluijter and Van Heuven (1996) investigated the effects of lexical stress on intensity, by controlling accent in the stressed syllables and the unstressed syllables in Dutch, a stress-accent language. They not only concluded that intensity (spectral balance) was greater in the stressed syllables than in the non-stressed syllables, but also that the role of spectral balance (as an indicator of intensity) was greater in the higher frequencies of the spectrum than in the lower ones in the stressed syllables. Therefore, the role of intensity is not as stable as duration.

The second debate surrounds the role of pitch. The question about the role of F0 focused on whether F0 is a solid cue for stress location. Beckman and Pierrehumbert (1986) found that duration and intensity were better predictors than pitch in perception of stress. Kochansky et al. (2005) found that there was no F0 role in the perception of stress-accent language, when they manipulated F0 in the real corpus data in which the prominent syllables and non-prominent syllables were marked. The most striking result was that many prominent syllables, as well as non-prominent syllables, had higher fundamental frequency. This supported the hypothesis that listeners did not use fundamental frequency as a cue for assigning stress. The role of F0 seems to be confounded with sentence level accents, and in the absence of accents (F0 at the sentence level), F0 may lose the role as a cue for stress.

The inconsistent results of F0 role are not only confirmed in perception studies, but also in production studies. Unlike the previous research in which Fry (1955, 1958) used a limited number of minimal pairs, Sereno and Jongman (1995) chose 16 disyllabic words from a large corpus in which the stress pattern was equally dominant, and participants read noun class and verb class words. They measured the duration, intensity, and average fundamental frequency, as acoustic information for stress location. They found that speakers were not only aware of the stress pattern, but also used acoustic information to differentiate reading of nouns and verbs. In the production, duration and intensity were strong cues for grammatical class, compared to fundamental frequency.

Moreover, the early research by Huss (1978) measured acoustic parameters of disyllabic minimal pairs in the nuclear position and post-nuclear position. He found that there was a difference between the stressed syllables and the unstressed syllables in terms of fundamental frequency in the nuclear position; however, this difference disappeared in the non-nuclear position, but the participants still could perceive a word as noun versus verb. This research teased apart the sentence intonation from lexical pitch. In strong position of F0, F0 role is limited in English.

The tentative role of pitch also appeared in languages other than English. Gordon (2004) tested eight native speakers in production of polysyllabic words in Chickasaw, a Native American language. He compared three stress levels in terms of F0, duration and intensity. He investigated whether the primary and secondary stressed syllables are acoustically different from each other, and from unstressed syllables.

Gordon (2004) found that the words with short vowels did not produce significant effects in fundamental frequency. But F0 was significant in distinguishing the primary stressed syllables from the secondary and the unstressed syllables, in syllables with long vowels. The increased duration is the cue for stress in Chickasaw. Moreover, the intensity was systematic, significantly differentiating the primary stress from the secondary and unstressed one. In terms of F0, Chickasaw used F0 in long vowels, but not in short vowels. The Chickasaw example suggested that the role of F0 depended on the vowel length.

Finally, the inconsistent results of F0 also included other languages, in which there was no F0 role at all for distinguishing the stressed syllables from the unstressed syllables. The Kuot language is an example that provided such information. Lindström & Remijsen (2005) examined the interaction of the F0 contour with lexical stress in Kuot language. By measuring F0, duration and vowel quality of the first two syllables, Lindstrom & Remijsen (2005) found that even when they were in different intonational positions, duration and vowel quality, not F0, provided the cues for the stressed syllables compared to the unstressed syllables. Non-significant F0 results showed that even though there were strong F0 changes, F0 was not as involved in distinguishing the stress location as other parameters (duration and vowel quality). Lindstrom & Remijsen (2005) concluded that Kuot is a language that uses duration and vowel quality to indicate the stress.

To summarize the relationship between acoustic correlates of duration, intensity and F0 to stress assignment, duration was the strongest cue for distinguishing the stressed syllables from the unstressed syllables, across languages that are considered as stress-accent languages. Intensity could be a cue for differentiating the stress location across languages. However, the role of intensity was not as consistent. It may be possible that the relationship between F0 and intensity restricts the role of intensity. This would occur when the falling F0 in stressed syllables might cause a decrease in intensity. The double function of F0 in lexical pitch and in sentential pitch affects the role it plays. Pitch has double functions that could be represented at the lexical level, as a lexical pitch, or at the sentential level, as an intonation pitch (Cutler, 1980). Even though pitch rises in stressed syllables, and rapidly drops in unstressed syllables, the pitch differences between stressed and unstressed syllables are minimized, or disappear in English, if the stress-contrasted words occur in question type intonation. In addition, if the stress-contrasted minimal pairs appear in the same position (non-focus position) in sentences, the pitch differences of the minimal pairs are too small to signal the stress (Ladd, 1996).

Whether F0 is used or not depends on how the sentential accent is teased apart from lexical accent. In the separated condition of stress and accent, F0 may not play a role in distinguishing the stressed syllables from the unstressed syllables. Even in the strongest condition, lexical stress does not rely on F0 to differentiate stress location. Lindstrom and Remijsen (2005) indicated that stress-accent languages use parameters other than F0. Therefore, the most common hierarchy of acoustic parameters is duration, intensity and then F0 in the production of stress in stress-accent language.

1.2.2 Vowel quality

In addition to duration, intensity and F0, vowel quality could be another parameter in stress-accent languages that provides a cue for stress assignment. Even though vowel quality also is influenced by many factors including focus, speech rate, formal speech or casual speech, dialects, syllable structure and position in speech, it has been investigated extensively as a cue for stress (Van Bergem, 1995; Fourakis, 1991; Fleming, 2005). In the stressed syllables, vowels are realized as full vowels, but in unstressed syllables, vowels are realized as the central vowels or schwa-like reduced vowels. The changes of vowel quality are called vowel reduction and measured by changes of first formant frequency (F1) and second formant frequency (F2) in the vowels (Delattre, 1969; Lindblom, 1963). In the same environment, unstressed vowels lose their quality and make a clear contrast against their stressed counterparts.

Different patterns of vowel reductions (vowel quality changes) were based on vowel spaces, and the distribution of front/back, or high/low vowels, in the vowel system in a language. Delattre (1969) examined the first two formant frequencies, of the stressed and unstressed vowels, that occurred in the medial positions in English (12 vowels), French (10 vowels), German (14 vowels), and Spanish (5 vowels). He found that the degree of vowel reduction was different in the four languages. English had the strongest vowel reduction, French and German had weaker reduction, and Spanish had the least.

The four languages were not only different in degree of vowel reduction, but also in the patterns of the vowel reductions. The vowel reduction in English had tongue centering, with the location of vowel reduction similar to the vowel /ʌ/. The vowel reduction in French was smaller than the one in English; it had tongue centering, with lip un-rounding, or lip un-spreading, based on the different vowels in the tongue positions. The vowel reduction in German was much smaller than English and slightly smaller than French. The acoustic pattern of vowel reduction was lip un-rounding or lip un-spreading, rather than tongue centering. The vowel reduction in Spanish was the smallest compared to the other languages. Here the acoustic pole of unstressed vowel attraction was in a medial position, slightly back to center, with no tongue fronting, which is different from the vowel reduction in the other languages. Therefore, even though all four languages are stress-accent languages, the degrees and location of vowel reductions are different as stress cues.

In the phonetic-based research of the stress-accent languages, phonetic (acoustic) vowel reduction is considered to be a parameter for providing a cue for stress assignment. Lindblom (1963) studied Swedish vowels in nonsense words, to examine vowel reduction in Swedish. He found that formant frequencies changed according to the decrease of vowel durations. The second formant frequency changed significantly.

In addition, according to Van Bergem (1993), acoustic vowel reduction indicates that "loss of vowel quality" in the spectrums is supposed to be fully expressed. He examined 'candy', 'canteen' and 'can' in accented versus unaccented position in Dutch. He measured the first two vowel formant frequencies, and spectral reduction of vowels, which are distanced from a full vowel to a reduced one. He found that spectral reduction was bigger in unstressed syllables than in stressed syllables, even though accented syllables had less spectral reduction than unaccented ones. In Dutch, vowel reduction occurs in unstressed and unaccented syllables, with the majority of vowels moving to schwa-like reduced vowels. Phonetic vowel reduction appears mostly in heavy stress-accent languages, such as Dutch, English, and Swedish (Bergem, 1995).

1.3 Stress patterns of Uyghur

1.3.1 The Uyghur language

The Uyghurs live in Xinjiang Uyghur Autonomous Region in Northwest China, with a population of over 9.65 million (Xinjiang Weiwuer, 2008). The following map shows where the Uyghurs live in China. In addition, there are many Uyghur groups outside of Xinjiang (e.g. Central Asia, Turkey, Australia, Europe and USA).

Figure 1 Xinjiang map (Central Intelligence Agency, 2011. China. Administration Divisions [map]. Scale 1:30,000,000. [Wash. DC: CIA]),

Uyghur, (ISO 639-3: uig), is Southeastern Turkic languages. Uyghur is an agglutinative language, rich in morphology and an OV typology. In Uyghur, being a Turkic lexical stress in Turkic language, typically stress occurs on the final syllable of a word (Muti, 2007), in polysyllabic words of Turkic origin. Therefore, Uyghur is known as a final stress language. However, sometimes stress is not in final position in Uyghur. For example, some negation suffixes and enclitics do not attract stress, and stress falls in non-final position (Muti, 2007: 43).

These exceptions usually occur in non-Turkic borrowings or in Turkic words that originated from different word classes. To date, there has been no systematic research, neither in phonology nor phonetics, on the word prosody pattern of Uyghur. Therefore, examining the stress pattern in Uyghur will provide new angles for investigating the typology of word prosody.

The accent pattern of Uyghur has not been extensively investigated. There is no systematic research to show whether Uyghur is a stress-accent language or a pitch-accent language. For Turkish, Levi (2005) used the same methods as Beckman (1986), and investigated the accent pattern in Turkish. She focused on inflection-level words instead of stem-level words. She came to the same conclusion that F0 was the only cue for distinguishing accented from unaccented syllables in Turkish. Her research on Turkish gives us clues about the possible accent pattern in Uyghur if we predict that Turkish and Uyghur should pattern the same. Both languages have similar accent patterns. If the accent pattern found in Turkish (or Japanese) is found in Uyghur, Uyghur should be considered a pitch-accent language.

1.3.2 Acoustic research on stress patterns in Turkic languages

For the majority of Turkic languages, acoustic research of stress pattern has not been widely investigated. The most extensive research has been done on Turkish. Stress pattern of Turkish was done by Levi (2005). By measuring F0, duration and intensity in near-minimal pairs produced by 10 female native speakers in Turkish, she found that only F0 could distinguish the stressed syllables from the unstressed syllables; therefore, she concluded that Turkish is a pitch-accent language as mentioned above. Even though, with absolute value comparison, F0, duration and intensity were significantly different in stressed syllables than in unstressed syllables, when she used Linear Discriminant Analysis, the order of importance was F0 (90% of data), intensity (70% of data), and duration (65% data) by single predictor. In step-wise LDA model, F0 was in the top position that can correctly classify over 99% of data for verbs and 96 % of data for nouns. For Turkish, Levi concluded that only F0 can distinguish the stressed syllables from the unstressed syllables.

However, in other Turkic languages (e.g. Chuvash), the acoustic research on stress patterns suggested that Chuvash has similar pattern as stress-accent languages, and mostly are not like pitch-accent languages such as Japanese. For example, Dobrovolsky (1999) investigated the stress pattern in Chuvash, a Turkic language. Instead of using minimal pairs, Dobrovolsky (1999) controlled the disyllabic words in four word-stress groups: full-full (first syllable is full vowel and second syllable is also full vowel), full-reduced, reduced-full, and reduced-reduced disyllables. Dobrovolsky (1999) found that words with the same stress placement were distinguished by duration ratios and total amplitudes in Chuvash. The duration ratios and total amplitudes were significant cues within each word-stress class except for the reduced-reduced group. For Chuvash, duration and total amplitude (intensity) are critical cues for assigning the stress location. Dobrovolsky also claimed that Chuvash stress may not always be in the final position. Based on Beckman’s (1986) distinction between stress-accent and pitch-accent languages, Chuvash should be a stress-accent language. Thus, research suggests that some Turkic languages may not be pitch-accent languages, in contrast to Turkish.

Among the few studies about the acoustic analysis of stress patterns in Turkic languages, two are on Uyghur. Liang and Zhang (2008) examined the stress pattern in Uyghur using 16 disyllabic words and 5 newly created non-words. The words were not contrasting in terms of stress, but the authors assumed the stress was on the final syllable. Ten native Uyghur speakers produced the words in isolation. Their stimuli had four syllable groups (CVCV, CVCCVC, CVCVC, and CVCCV syllable structures). Liang and Zhang measured the fundamental frequency, duration and peak intensity of both syllables in each word. They compared the first syllables to the second syllables with absolute values for each acoustic parameter and concluded that only duration provided a stress cue because the duration of all second syllables was longer than the first syllables both in real words and newly created non-words. In all the productions, second syllables (which carried stress) had longer duration than the first syllables, which were not stressed. Therefore they concluded that the stressed syllables had longer duration than unstressed syllables. The accented syllables did not have higher fundamental frequency and greater intensity.

In a second study, Jiang et al (2010) used disyllabic and trisyllabic words to examine the accent pattern in Uyghur using 12 native Uyghur speakers. They also assumed that stress was in final position of a word. They measured average F0, duration, and total intensity for each syllable, and they also used ratios of the second syllable to the first syllables. Even though they did not provide any statistics, they found that second syllables were longer in duration or greater in intensity than the first syllables because of the ratios were bigger than 1. Therefore, they concluded that only duration and intensity effectively distinguished the stressed syllables from the unstressed syllables. However, neither of these studies used minimal pairs that completely control the environment in which stressed and unstressed vowels occurred and both studies always assumed that stress was in final position. In addition, both studies measured the acoustic parameters based on entire syllables and phonetic content was not controlled. Therefore, the results could be due to final lengthening because of isolated production and could be influenced by different phonetic contents. We systematically examined the minimal pairs and disyllabic nouns contrasted on the first syllables.

1.4 Stress patterns in non-native speakers

“Foreign accents” often reflects the transfer of some components from the L1 (native language) to the L2 (target language). In other words, the inappropriate transfer from the L1 to the L2 causes the foreign accent in L2 speakers. Foreign accent is influenced by intonation, segmentation and prosody of the native language (L1), when people use a second language (L2).

Stress acquisition is one of the factors that influence the foreign accents. Under this assumption, stress patterns can be learned by non-native speakers via an inter-language stage (Erdmann, 1973). By using known and unknown words in English, Erdmann (1973) found that some rule-based stress patterns of English were acquired by German native speakers. In his research, words with initial stress in English were produced by German learners of English neither like English nor like German. For example, the adjectival suffixes in German are stress-fixed while the adjectival suffixes are non-stress attracting in English. Stress patterns of German-English bilinguals were different from Native English and German native speakers. German learners of English used stress on the penultimate in disyllabic words, and they put stress on the antepenultimate position for trisyllabic words in English. Erdmann (1973) focused on phonological analysis not phonetic based research. His idea supports that learners can acquire stress via an intermediate stage.

However, Zuraiq and Sereno (2007) investigated the acquisition of stress patterns in English by Jordanian Arabic speakers. Jordanian Arabic speakers learned English trochaic and iambic stress patterns in nouns versus verbs in English by using similar parameters including fundamental frequency, duration and intensity as cues similar to English native speakers. Jordanian Arabic speakers did not use vowel reduction as English native speakers did. Even though Jordanian native speakers also used fundamental frequency, duration and intensity in Jordanian Arabic minimal pairs, they also used similar parameters to distinguish stress in English except vowel reduction (Zuraiq, 2005). We cannot tell whether the acquisition of stress was due to transfer or independent learning process of stress.

Other studies focus on what component of the L1 transfers into the L2. Non-native speakers may transfer the L1 stress patterns to the L2. If the L1 uses some phonological and phonetic parameters in the stress pattern, the L2 learners may use the same parameters in their L2 production. For example, if a language has trochaic stress in disyllabic words in the L1, then the L2 learners may use a trochaic stress pattern in the L2, even though there is no trochaic stress pattern in the L2. Archibald (1992, 1993, and 1998) had extensively investigated the L2 learners acquiring stress in terms of phonological properties. In his research, he examined English stress acquisition by Polish, Hungarian, and Spanish learners of English. He found the following patterns. (1) In Spanish, there was no clear transferring, because of similarity of Spanish and English stress patterns, unless there were extrametricality markings in Spanish, which did transfer into their L2 production of English. (2) The transferring of the L1 Polish patterns to English was significantly shown by the stressed penultimate in English, which is a common characteristic of Polish, but not English. (3) Hungarian has initial stress and long vowels that attract stress. Both of these phonological characteristics of Hungarian were transferred into English production by Hungarian L1 speakers. It seems therefore, that in acquiring a stress pattern in the L2, the dissimilar components of the L1 transfer into the L2 production. Archibald (1997) also examined the process by which the L1 speakers of non-stress accent languages such as Japanese and Chinese, in which the L2 learners of English have a non-stress accent language background, acquired the stress-accent pattern of English. The results suggested that L2 learners produced more native-like stress pattern in English, rather than transferring the L1 components into L2. Therefore, in acquiring the L2 stress pattern, there were at least two components including the transfer of the L1 stress pattern to L2, and the learning of the rules of the L2 patterns independently.

However, the linguistic transfer or the learning of the L2 stress pattern is influenced not only by the similarities and dissimilarities between the L1 and the L2 (Archibald, 1992, 1993, and 1997), but also by the acquisition age of the L2 (Guion et al, 2004; Pater, 1997). First, Guion et al. (2004) examined English native speakers, early Spanish-English bilinguals and late Spanish-English bilinguals using disyllabic non-words. They manipulated (1) the syllable structures, which were more like English syllable structures including four syllable structures (CVVCVCC, CVCVCC, CVCVC and CVCVVC); (2) lexical class, which is simulated by noun versus verb disyllabic words that had initial versus final stress; (3) phonological similarity, where non-words sound more like English real words. In the production of the non-words by three different groups (Experiment 1), they found that both native English speakers and early Spanish-English bilinguals produced a similar pattern in which they had more initial stress on non-words in a noun frame than in a verb frame, and they had sensitivity for long vowels that were stressed easily. But the late Spanish-English bilinguals had weak and somewhat different lexical effects and syllable structure effects. Late Spanish-English bilinguals did not use the effects of lexical class on stress assignment in the production and the late bilinguals had more initial stress, which was overgeneralized from the distribution of stress placement in English.

From Guion et al's study, we learned that even though both early (acquisition age 2.5-5 years old) and late (acquisition age 15-33 years old) Spanish-English bilinguals had similar proficiency in English, the production patterns were different, in which the early Spanish-English speakers were similar to native English speakers, while the late Spanish-English speakers were slightly different. The age of acquisition could be one factor that influenced the acquisition of stress patterns. Therefore, the nature of linguistic transfer in the stress pattern domain varies according to the characteristics of the L1 and the L2, and the dominance of the L1 over the L2. The transfer may be one way in which the L1 influences the L2, and/or vice-versa. Therefore, it is important also to investigate not only L1 stress patterns, but also to investigate the acquisition of the L2 stress patterns by non-native speakers.

In contrast to the above studies about the L2 stress pattern acquisition, which focused on phonological aspects of acquisition, the phonetic properties of stress parameters of the L1 also transfer to the L2 or can be learned by non-native speakers. Mennen (2006) indicated that the prosody transfers via phonological and phonetic means. Acoustic lexical stress has been investigated not only among native speakers, but also among learners (Zuraiq and Sereno, 2007; Lai, 2008). Zuraiq and Sereno (2007) examined the L2 acquisition of English stress by Arabic native speakers whose second language is English. By measuring fundamental frequency, duration, intensity and vowel reduction in eight English minimal pairs, they found that English native speakers consistently used four parameters. However, Arabic speakers only used duration and intensity in a similar way to native English speakers, and they did not make similar use of vowel reduction. In terms of fundamental frequency, Arabic speakers used F0 to a larger extent compared to native English speakers.

In order to know how their native language influences stress assignment in their second language, Zuraiq and Sereno (2007) and Zuraiq (2005) also compared six minimal pairs in Arabic. They found that Arabic speakers consistently used fundamental frequency, duration and intensity, but not vowel reduction, in assigning the stress in their native language. They concluded that Arabic speakers are like English native speakers in using duration and intensity. However, Arabic speakers overused fundamental frequency in English and did not use vowel reduction. The advantages of this research were as follows: (1) it included the stress patterns of English and Arabic speakers in their own native language and then examined the stress pattern acquisition by non-native speakers. (2) proficiency level was manipulated as beginner and advanced level of English speakers. High proficiency level learners were more similar to native speakers, and the lower proficiency learners had different patterns from native speakers. They found that proficiency was one factor that influenced the acquisition of stress patterns.

In addition, Lai (2008) investigated the acquisition of English stress patterns by Mandarin Chinese speakers. She also examined proficiency levels with beginning and advanced level of English. By measuring maximum F0, mean F0, duration and intensity, she found that beginning levels and advanced levels of English learners were similar to native English speakers in terms of using acoustic correlates to stress for nouns. But for verbs, the learners used amplified duration effects and minimized F0 effects. English learners of Mandarin Chinese used similar patterns as English native speakers, but they used consistent F0 cues in verbs due to the influence of Mandarin Chinese. Therefore, F0 acoustic parameters were transferred from the L1 to the L2.

However, Lai did not find differences between beginner and advanced levels of English learners. Zhang et al. (2008) also investigated the acquisition of English stress patterns by Mandarin Chinese speakers. Even though native speakers and English learners were similar in terms of using duration and intensity, the learners overused F0 in English. In the acquisition of English stress pattern by Jordanian Arabic speakers, Zuraiq and Sereno (2007) found that these Arabic speakers also overused F0 effects in producing English minimal pairs that contrasted in stress location.

In sum, previous research has focused on the acquisition of English stress patterns by non-native speakers. There is little research that focused on the L2 acquisition of stress patterns in languages other than English. One of the few, Özçelik (2012) examined acquisition of Turkish stress pattern by French and English native speakers. According to that study, English speakers could not learn Turkish stress regardless of proficiency; in contrast, French learners of Turkish could learn Turkish stress, because French has similar stress patterns to Turkish. Özçelik considers both Turkish and French to be footless and the role of F0 to be at the intonational level rather than the lexical level. His research focused on phonological aspects, not acoustic analysis of the parameters. His results suggested that the similarity of the L1 and the L2 can cause the transfer rather than the independent learning of stress.

From the previous research, we might consider the language similarity and dissimilarity between the L1 and the L2, as well as proficiency in the L2 as important factors. If we focus on the acoustic research on stress patterns by non-native speakers, we find that non-native speakers learned the stress pattern in English (Zuraiq & Sereno, 2007, Zhang et al, 2008, Lai, 2008) regardless of native language type. However, English is the target language in this research. Our focus is on Uyghur, a less commonly taught language, as the L2 language. The result might provide different insight into the acquisition of stress patterns across languages.

Rationale for studying Uyghur stress pattern

First, acoustic methods will provide reliable data to investigate stress. Secondly, the diversity of the research on accent (stress) patterns of languages will help us to address the nature of stress patterns among languages. Thirdly, acoustic research is an important tool when the accent or stress patterns are not stable or not agreed on by native speakers. Uyghur is a case in which different native speakers have different intuitions about accent patterns. Finally, the results of acoustic research could be employed in developing a phonological analysis of stress in Uyghur.

1.5 The present study

Lexical stress patterns in languages have been investigated for more than half a century (Fry, 1955, 1958; Beckman, 1986; Sereno and Jongman, 1995; Sluijter & Van Heuven,1996; Van Der Mark, 2003; Levi, 2005;). The main concern focuses on which acoustic paramters provide cues for assigning stress. English is the most investigated languge in which fundemnetal frequency, duration, and intensity provide phonetic correlates of English lexical stress. On the other hand, Beckman (1986), who acoustically distinguished pitch-accent languages from stress-accent languages, proposed that languages use multi-dimensions or single dimensions such as pitch is the best predictor to differentiate the pitch-accent languages (Japanese) from the stress-accent languages (English). Her identification of the differences between stress-accent languages and pitch-accent languages focus on the phonetic cues (Beckman, 1986). Stress-accent languages draw information from pitch, intensity, and duration as cues to distinguish the stressed syllables from non-stressed syllables. On the other hand, in pitch-accent languages, pitch is the primary cue to distinguish the accented syllables from the unaccented syllables. In other words, stress-accent languages use multiple dimensions including pitch, duration, and intensity as cues for distinguishing stressed from non- stressed syllables while pitch-accent languages use primarily one dimension (pitch) for differentiating the accented syllables from the unaccented syllables. The present research will examine the acoustic parameters of stress in Uyghur. The same acoustic parameters including fundamental frequency, duration and intensity for distinguishing stress location will be examined. If Uyghur uses all three of these parameters (fundamental frequency, duration and intensity) or some combination of these parameters, we could conclude that it is a stress-accent language. If Uyghur only focuses on fundamental frequency but no other parameters, it can be considered as pitch-accent language like Japanese.

Levi (2005) examined the accent pattern in Turkish by measuring fundamental frequency, intensity and duration of accented versus unaccented syllables. Levi (2005) concluded that Turkish is a pitch-accent language because only fundamental frequency differentiated the accented syllables from unaccented syllables. This research on Turkish gives us possible clues about the accent pattern in Uyghur, which is a branch of Turkic, rich in morphology, and an agglutinative language. Uyghur may be considered a pitch-accent language, if the acoustic correlates of stress that are found in Turkish are also found in Uyghur. Unlike Turkish, Uyghur may use more parameters other than pitch; in that case, Uyghur should be categorized a stress-accent language. The basic research question is which parameters provide the cues for stress locations in Uyghur. The parameters that will be examined are fundamental frequency, duration, and intensity.

1.5.1 Experiment 1 Acoustic study of minimal pairs in Uyghur

Experiment 1 examined a set of minimal pairs that contrasted in terms of the location of the stressed syllables. In this experiment, five native Uyghur speakers produced six minimal pairs which contrast in terms of stress placement, including six initial stress and six final stress words. We measured three acoustic parameters: average fundamental frequency, duration, and average intensity. For the analyses, we used ratios of these values as Beckman (1986) indicated. We predict that if Uyghur is a stress-accent language, fundamental frequency, duration and intensity may be strongly correlated to stress location. We would then conclude that Uyghur is a stress-accent language. On the other hand, Uyghur could be a pitch-accent language. In that case Uyghur may only use F0 as a stress cue (cf. Levi, 2005). The result of this experiment helps us to categorize Uyghur as a pitch-accent or stress-accent language based on its acoustic properties.

1.5.2 Experiment 2 Acoustic study of disyllabic nous that contrast in the first syllable

In Experiment 2, instead of using minimal pairs, we examined disyllabic words which share the same syllable in the initial position. In these pairs, the first syllable in one member of the disyllabic words is stressed, and the first syllable is not stressed (second syllable is stressed) in the other member of the disyllabic pair. For example, in *DAka* 'gauze' versus *daLA* 'plain,' the first syllable 'da' of the first disyllabic word (DAka) is stressed, and the first syllable 'da' in the second disyllabic word (daLA) is not stressed. We can compare the values of both stressed and unstressed syllables in the disyllabic words which have identical syllabic context. Using stimuli that contrast in stress location but are not minimal pairs will allow a more representative set of stimuli to be examined. Comparing identical syllables across disyllabic words will provide an excellent control of context.

Experiment 2 manipulated the influence of syllable structure (e.g. CV vs. CVC) in stressed syllables. There is little research on whether syllable weight matters in Uyghur. Syllable weight might play a role in assigning stress in Uyghur. In order to control the possible effect of syllable weight in this experiment, we will use two syllabic structures, CV versus CVC syllables, which are the most frequent syllable structures in Uyghur (Saimaiti and Feng, 2008). Finally, Uyghur is a case in which different native speakers seem to have different intuitions about accent patterns. Therefore, we used a pre-test with a number of disyllabic words (75 pairs across CV and CVC structures) and had native speakers indicate where the stress was located. This allowed us to seperate the highly consistent and less consistent stimuli among them, since consistency may influence the acoustic parameters. We chose six pairs for each CV-consistent group, CVC-consistent and CV-inconsistent group, respectively.

We examined whether the acoustic parameters in Experiment 1 could generalize to a new set of stimuli in Experiment 2. We predict that Experiment 1 and Experiment 2 may have similar results in terms of using acoustic cues.

1.5.3 Experiment 3 The interaction between lexical stress and sentential intonation

Experiment 3 examined the interaction between lexical stress and sentential intonation. While fundamental frequency, duration, and intensity have been established as cues for assigning stress location in language, the relative importance of these parameters may change based on the positions of target words in the sentences. The position of target words in different contexts, especially, for example, interrogative sentences, will shed light on the interaction between the context and lexical stress and the hierarchy of involvement of the acoustic parameters in producing lexical stress. In a strong position of F0 (e.g. interrogative sentences), F0 as a cue to stress may be enhanced or diminished. To examine these, we used CV-consistent group stimuli in declarative assertion (DA) and declarative question (DQ) type of sentences. The target words occurred in the final position in both sentence types. We measured the same parameters that we used in Experiment 1 and Experiment 2.

We predicted that F0 might be used in the DQ condition due to the declarative question sentential intonation and not due to lexical stress. The main motivation of Experiment 3 is to clarify if F0 cues are the result of sentential intonation rather than lexical stress.

1.5.4 Experiment 4 The acquisition of stress pattern by non-native speakers

Experiment 4 examined the acquisition of stress patterns by non-native speakers. Recently, the use of lexical stress has been investigated not only in native speakers, but also among learners (Zuraiq and Sereno, 2007; Lai, 2008). What acoustic cues do English native speakers use when they produce Uyghur? In order to answer this question, we examined Uyghur leaners, whose native language is English, producing minimal pairs, and disyllabic words with contrasting stress. We measured the same parameters in previous experiments. We used minimal pairs (similar to Experiment 1) in Experiment 4a and disyllabic noun pairs (similar to Experiment 2) in Experiment 4b in order to compare these results against those of native Uyghur speakers. Five non-native Uyghur speakers (highly advanced) were recorded. All of the learners had at least 4 semesters and are still working in Uyghur culture or language.

We predicted that learners may use similar parameters as Uyghur native speakers do. If they are influenced by their native language (English), they might use a number of parameters, indicating duration, intensity and especially fundamental frequency.

Chapter 2 Experiment 1 An Acoustic Study of Minimal Pairs in Uyghur

2.1 *Previous research and research questions*

2.1.1 Previous research

A majority of research (Fry, 1955, 1958; Beckman, 1986, Van Der Mark, 2003; Sluijter and Van Heuven, 1996a, 1996b; etc.) focused on minimal pairs when examining stress patterns. Minimal pairs were contrasted in terms of stress location, and were segmentally equal to each other. In the acoustic analysis of the stress pattern in language, fundamental frequency, duration and intensity were considered as relevant acoustic parameters.

Two studies were closely related to our research. Beckman (1986) compared the accent patterns in Japanese and English, and concluded that English was a stress-accent language and Japanese was a pitch-accent language. Levi (2005) investigated the accent pattern in Turkish, and found that Turkish was a pitch-accent language rather than a stress-accent language.

Beckman (1986) compared two different types of accent languages using acoustic analysis by measuring fundamental frequency, duration and intensity in the production of disyllabic words in Japanese and English. She used six minimal pairs in Japanese, part of a larger corpus used in an experiment on segment duration. For English, she also used the five minimal pairs that were used by Fry (1955, 1958). In Japanese, the minimal pairs were fixed in the same carrier sentences. However, in English, subjects read the target words in two different context sentences. For example, if the target word was *permit*, there were two context sentences for 'permit'.

> *In order to park here you need a permit.* (with *permit* as a noun, and the stress is on the first syllable, PERmit)
> *Would you permit it?* (with *permit* as a verb and the stress is on the second syllable, perMIT)

The stimuli were produced by one female and four male Japanese speakers for the Japanese minimal pairs and by two male and two female native English speakers for the English minimal pairs.

Beckman also measured the peak fundamental frequency (peak F0), duration and amplitude (peak amplitude, total amplitude and average amplitude) for the stressed and the unstressed syllables. Instead of comparing all of the parameters directly in stressed and unstressed syllables, she used ratios of all the parameters in stressed and unstressed syllables in the same word.

The ratio of F0 values (in semitones) was defined as:

$$\text{Semitone difference} = 17.31 \ln[\text{Hz}(S2)/\text{Hz}(S1)]$$

Here, Hz (S1) was the measured peak fundamental frequency value of the first syllable and Hz (S2) was the measured peak fundamental frequency of the second syllable. A negative value indicated a falling F0, and a positive value indicated a rising F0. She also defined the three different measurements for intensity:

$$\text{Peak amplitude ratio} = dB_{peak}(S2) - dB_{peak}(S1)$$

Because the intensity was measured in decibels, the ratio was calculated by subtracting first syllable intensity from second syllable intensity. In peak amplitude ratio, $dB_{peak}(S1)$ is the peak amplitude value of the first syllable and $dB_{peak}(S2)$ is the peak amplitude of the second syllable. The total amplitude ratio = $dB_{total}(S2) - dB_{total}(S1)$, was calculated by adding up measurements taken at every 10-ms intervals on the vowel.

$$\text{Average amplitude ratio} = dB_{average}(S2) - dB_{average}(S1)$$

The average amplitude was derived from second measurement in which total amplitude was divided by duration and multiplied by 10. A positive value indicated a rising intensity and a negative value indicated a falling intensity.

Log duration ratio = ln[ms(S2)/ms(S1)], was calculated from the logarithm of the quotient of duration of two syllables. A positive value indicated that duration was increased from the first syllable to the second syllable, and a negative value indicated that duration is decreased from the first syllable to the second syllable.

Beckman's data showed that each token in stressed and non-stressed syllables was significantly different in F0 ratios both in Japanese and English. In other words, F0 ratio can distinguish stressed syllables from unstressed syllables. Moreover, in terms of the peak amplitude, all English minimal pairs showed significant differences in peak amplitude and average amplitude values except for 'permit,' and, 'contrast'; however, Japanese minimal pairs were not distinguished by peak amplitude or average amplitude. This suggested that the peak amplitude and average amplitude differentiated the stress pattern in English, but not in Japanese. Furthermore, the duration and total amplitude ratios differentiated the stress pattern in English except where mentioned for the two minimal pairs, but not in Japanese. Therefore, the accentual pattern had considerable effects on both the duration and total amplitude patterns in English, but not in Japanese.

To distinguish the pitch-accent and stress accent languages, Beckman concluded that one criterion to differentiate pitch-accent and stress-accent languages was that stress-accent languages used F0, duration and amplitude as cues for stress-accent patterns, and Japanese, as a pitch-accent language, only used F0 as a cue.

Whereas the previously described Beckman's study focused on the word-level stress pattern, Levi's (2005) Turkish study also used acoustic analysis to investigate the stress pattern at the inflectional level. Regarding the Turkish language, some studies of Turkish (e.g., Sezer, 1983, cited from Levi, 2005) claimed that Turkish was a stress-accent language while others (e.g., Underhill, 1986, cited from Levi, 2005) favored the opinion that Turkish is a pitch-accent language. Like Beckman (1986), in her study, Levi (2005) tried to classify Turkish as a pitch-accent or stress-accent language by comparing acoustic differences between the stressed and the non-stressed syllables.

In Turkish, the default stress is on the final syllable, but there are pre-stressing suffixes such as *-ma/me* (negation markers) or *-le* (post position) for which the stress is located before the pre-stressing suffix. It is hard to find word (stem)-level minimal pairs such as in English or Japanese; therefore, Levi (2005) used two sets of word pairs at the inflectional level: (1) a noun group with10 disyllabic noun stems, with one of two suffixes */-de/*, which is stress-able, or */-le/* which is pre-stressing. (2) a verb group with 10 verbal stems with one of two suffixes, either the verb plus gerundial */-mAk/* (which has the final stress), or the verb negation suffix, *-mA* (negation infixes) followed by gerundial, was added before the verb endings with */-mAk/*, and the negation blocked the stress moving to right forward; that caused that the stress was on the original syllables. For example, the verb, *bar* 'go' with *barmak* 'going' versus *BARmamak* 'not going' made near minimal pairs in terms of stress location, Therefore, the nouns and verbs with different suffixes or infixes created near-minimal pairs.

Levi (2005) put all target words in a fixed sentence. All target words were produced in three random orders by seven female native Turkish speakers. She focused on three measurements: F0, vowel duration and intensity. F0 was measured by excluding the initial and final 15ms and obtaining the peak F0 value in the reminder of the target vowel. For words beginning with oral stops, vowel duration was measured in two ways: ‘with VOT’ in which the release of the closure was the beginning of the vowel; and ‘without VOT’, in which the beginning of the vowel was calculated at the point of the first voicing pulse after release burst. In all other cases, the beginning of the vowel was the point in which there was a clear change in the waveform. Lastly, the peak intensity for each vowel in disyllables was measured. If the preceding consonant was a stop, peak intensity was recorded from the ‘without VOT’ segment measure. Levi (2005) used ratios like Beckman (1986) for both noun and verb pairs in which the lexically stressed and unstressed syllables were compared.

Levi (2005) found that in terms of F0, verbs with the lexical stress on the final syllable showed a rise in F0 from the initial to the final syllable, while verbs with an initial stress showed a drop in F0. Stressed syllables had higher F0s than unstressed syllables. In addition in distribution of F0 ratios, initial stressed words and final stressed words had a clear difference. In other words, there was no overlap between the initial stressed words and final stressed words in terms of F0 ratio distribution. In terms of intensity, verbs with the lexical stress on the final syllable showed a slight rise in intensity from the initial to the final syllables, while verbs with an initial stress showed a drop in intensity. Stressed syllables had higher intensity than unstressed syllables. However, there were big overlaps between the final and initial stressed words in terms of the distribution of the intensity ratios. In terms of duration, duration was calculated ‘with VOT ‘and ‘without VOT.’ For both ‘with VOT’ and ‘without VOT’, durations in stressed syllables were longer than the non-stressed syllables. However, the ratio distribution for initial and final stressed words overlapped with each other. The results in nouns were the same as verbs for all measurements.

Unlike Beckman (1986), Levi (2005) used Linear Discriminant Analysis (LDA) to evaluate classification accuracy that can be based on different parameters. Using a step-wise analysis, she also found that duration alone could not accurately classify the stress in both verbs and nouns; intensity increased classification from 93.5% to 96% in nouns, but in verbs, intensity could not correctly classified the stressed and non-stressed words. Only F0 correctly classify the nouns and verbs. Therefore, F0 was clearly the most important cue for stress location. Levi (2005) concluded that Turkish was a pitch-accent language.

The conclusions suggesting that fundamental frequency (F0) was a critical cue for differentiating the stressed syllables from the unstressed syllables was similar for Beckman (1986) and Levi (2005). When we compare two languages in terms of acoustic patterns, it is clear that stress-accent languages use multiple parameters in order to distinguish the stressed syllables from the unstressed syllables; however the pitch-accent languages use only F0 as the parameter to distinguish the stressed syllables from the unstressed syllables. Beckman (1986) focused on stem-level words. On the other hand, Levi (2005) used the same methods as Beckman (1986), and focused on inflectional-level words instead of stem-level words. She came to the similar conclusion that like Japanese, F0 was the only cue for distinguishing the stressed syllables from the unstressed syllables in Turkish.

2.1.2 Research questions and hypotheses

In this project, the stress pattern of Uyghur was investigated. Uyghur is a branch of Turkic languages. Levi's (2005) research on Turkish gives us clues about the stress pattern in Uyghur, because Uyghur belongs to the same family as Turkish and shares common characteristics with Turkish as an agglutinative language. If the stress pattern found in Turkish (or Japanese) is found in Uyghur, Uyghur can be considered a pitch-accent language. If Uyghur uses more than one parameter, it can be classified as a stress-accent language.

Based on the previous research, the research question was proposed as follows:

What acoustic parameters are strongly correlated to stress location in Uyghur?

In order to answer this question, two different hypotheses about the stress pattern of Uyghur were tested.

- If Uyghur is a stress- accent language, native Uyghur speakers will use fundamental frequency, duration and intensity to differentiate the stressed syllables from the unstressed syllables in the same manner as English native speakers.
- If Uyghur is a pitch-accent language, native Uyghur speakers will only use fundamental frequency for differentiating the stressed syllables from the unstressed syllables in the same manner as Japanese or Turkish speakers.

2.2 Methods

2.2.1 Stimuli

Minimal pairs are the best way to test the stressed versus unstressed syllables acoustically (Beckman, 1986; Levi, 2005; Fry, 1958), because they keep the phonemes and number of syllables the same for each member of the minimal pair. Therefore, stress was tested in conditions in which all aspects that add variability were controlled as much as possible and the remaining effects were due only to differences in stress.

In some minimal pairs in Uyghur, the stress is located in different positions, and thus the meaning is differentiated by the location of stress. For example, consider ‘Acha’ (elder sister) versus ‘aCHA’ (branching). In the first word of the minimal pair, the stress is located on the first syllable, and in the second word, it is located on the second syllable. A total six minimal pairs like this was used for this study as shown in Table 4.

Table 4 The Minimal Pairs in Uyghur Used for Experiment 1

#	Initial stress	English gloss	Final stress	English gloss
1	Acha	elder sister	aCHA	branching
2	Ara	fork	aRA	between
3	TÖshük	kitchen	töSHÜK	hole
4	BAla	child	baLA	disaster
5	CHAtaq	bad branch of tree	chaTAQ	problem
6	PAchaq	leg	paCHAQ	piece

(Capitalized syllable has stress)

Two syllable words were compared because they make clear contrasts between the stressed versus the unstressed syllables in minimal pairs, and controlling of the stress locations is easier than if polysyllabic words are used.

In addition, in order to cue the intended stress pattern, the stimuli were presented to the speakers in appropriate context sentences. For each token, speakers read the context sentences with the target words and then read a fixed sentence including the target words. The fixed sentences that include target words are identical across both minimal pairs. Finally, speakers read the target words in isolation. For example:

The three recording conditions for 'Acha'(elder sister) were:

(1a) Context sentence

Qeshqer-de öz-i-din chong qiz qérindash-i-ni Acha de-y-du.
Kashger-LOC[1] self-POSS-ABL big female sibling-POSS-ACC elder.sister say- PRE. 3rd pl.
'In Kashger, people called the elder female sibling as 'elder sister.'

[1] LOC is locative case; POSS is possessive case; ACC is accusative case; PRE is present tense; sg is singular, PL is plural; SIM/B represents simultaneous/ because; 'é' corresponds to the vowel /e/, and 'e' corresponds to /æ/ or /ɛ/ in IPA;

(2a) Fixed sentence

Men hazir ‘Acha’ de-y-men.
I now ‘elder.sister’ say-PRE.1^{st} sg
‘Now I say ‘elder sister’.

(3a) Isolation

Acha
‘Elder sister’

Three recording conditions for ‘aCHA’ (branching)

(1b) Context sentence

Bu yaghach aCHA bol-ghachqa bala-lar u-ni oyna-di.
This wood branching be-SIM/B child-PL it-ACC play-PAST.3^{rd}sg
‘This wood is branching, so kids use it.’

(2b) Fixed sentence

Men hazir ‘aCHA’ deymen.
I now ‘branching’ say-PRE.1^{st} sg
‘Now I say ‘branching.’

(3b) Isolation

aCHA
‘branching’

The context sentences provided a semantic background for each target word. The fixed sentences and isolation were the same for both members of the minimal pair. We used the fixed sentence context for all analyses.

2.2.2 Participants

Five native Uyghur speakers, two males and three females, living in Kansas participated in the study. The age range was from 24 to 70 years old. None of them had speech or hearing disorders. All speakers were native Uyghur speakers. All were also fluent in Mandarin Chinese and all of them had some knowledge of English, although two speakers had very little knowledge of English.

Each participant's speech sounds were recorded with a Marantz PMD 671 solid-state recorder. The speakers were recorded in an anechoic chamber via an Electro-Voice ND767a microphone at the University of Kansas.

2.2.3 Procedure

Each subject was recorded individually. Subjects were instructed to read the sentences that included context sentences followed by the fixed sentences, followed by the isolated target words for each member of a minimal pair. The minimal word pairs were presented in a random order three times, yielding 18 tokens for each Uyghur word type. The recordings were digitized at a sampling rate of 22.05 kHz. The total was 540 (5 speakers ×6 pairs × 2 words × 3 repetitions × 3 contexts) tokens for the three conditions (context, fixed and isolation).

In the current study, we only analyzed the stimuli from the fixed context sentence. Therefore, a total of 36 tokens were analyzed for each speaker. In sum, there were a total of 180 tokens (5 speakers × 6 pairs × 2 words × 3 repetitions) in this analysis.

2.2.4 Measurements

Three acoustic measurements were taken for each target word using Praat software (Boersma & Weenink, 2011): fundamental frequency (F0), duration, and intensity.

The first measurement was the fundamental frequency (F0). Starting with onset of the vowel, F0 was measured at 10ms intervals. The first 20ms and last 20ms of data removed from the data and the average F0 was calculated across these intervals, which was 60ms of data from the center of the vowel. We expected to find a higher average F0 in stressed syllables.

The second measurement was duration. Vowel duration was measured from wide-band spectrograms. The beginning of vowel was the point at which a clear first formant frequency (F1) changes (onset of F1) and the ending of the vowel is the point where a clear second formant frequency changes (offset of F2). A longer duration was expected from the stressed syllables.

The third measurement was intensity: The average intensity (in dB) was obtained from Praat based on the entire vowel. If intensity was related to the stress pattern in Uyghur, a higher average intensity would be found in the stressed syllables.

For minimal pairs, the lexically stressed and unstressed syllables were compared. The ratio comparison was used in which Beckman's (1986) formulas were used.

The ratio was calculated for average F0.

$$Semitone\ differences = 17.31 \ln \frac{[Hz(\sigma 2)]}{[Hz(\sigma 1)]}$$

In this formula, $\sigma 2$ is the second syllable and $\sigma 1$ is the first syllable in each target word. Hz ($\sigma 2$) indicates the measured fundamental frequency value of the second syllable while Hz ($\sigma 1$) indicates the measured fundamental frequency value of the first syllable. A negative value indicates a falling F0 pattern and a positive value indicates a rising F0 pattern.

The duration ratio was calculated as follows:

$$Log\ duration\ ratio = \ln \frac{[ms(\sigma 2)]}{[ms(\sigma 1)]}$$

The intensity ratio was calculated as follows

$$\text{Average intensity ratio} = \text{dB}(\sigma 2) - \text{dB}(\sigma 1)$$

2.2.5 Statistics

For statistics, Repeated Measures two-way ANOVAs were used for the ratio value measurement by the subject analysis and the item analysis. Separate ANOVAs were conducted for the average F0, duration and the average intensity. In addition, we used a paired t-test for the average F0, average intensity and duration ratios comparisons.

The average F0, average intensity as well as duration were calculated within words and across words. The two pitch tracks and spectrograms in Figure 2 and Figure 3 show that the pitch and intensity contours change for the minimal pair *Acha* (elder sister) vs. *aCHA* (branching). For example, in ‘*acha*’ as an ‘elder sister’ the vowel ‘*a*’ in the first syllable and ‘a’ in the second syllable in the ‘*cha*’ were measured. The minimal pairs ‘*Acha*’ (elder sister) vs.’*aCHA*’ (crossed) were compared in Figure 2 and Figure 3 below.

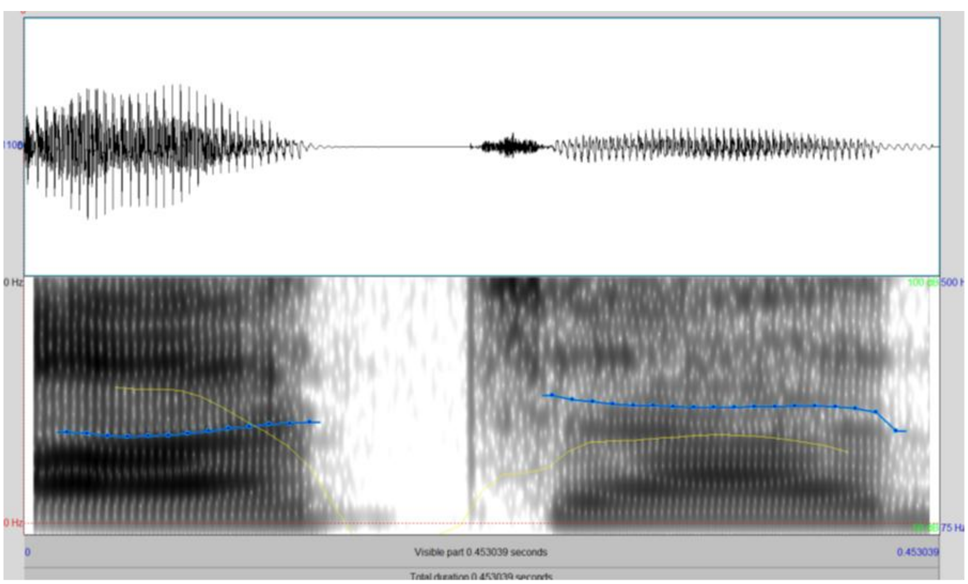

Figure 2 Waveform, spectrogram, pitch track and intensity contour of ‘Acha’ (elder sister) for speaker 1. The pitch range is 500Hz.

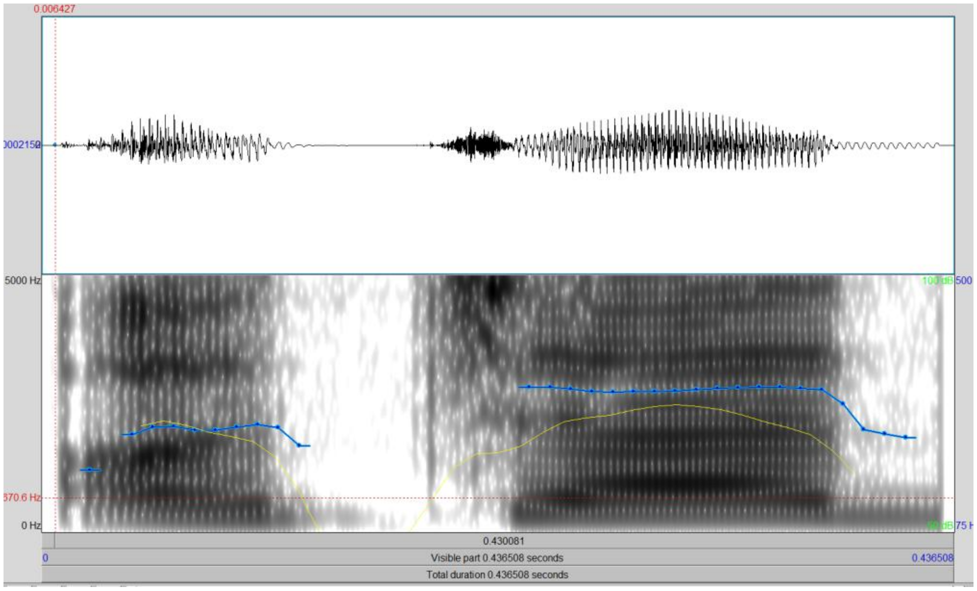

Figure 3 Waveform, spectrogram, pitch track and intensity contour of 'aCHA' (branching) for speaker 1. The pitch range is 500Hz.

Figure 2 shows that in initial stressed words, the pitch contour (blue line) had a slight rise but the intensity contour (yellow line) had a slight drop between syllable 1 and 2. However, in the final stressed word (Fig. 3), even though the pitch contour had a big rise, the intensity contour had also a rise for syllable 1 and 2. In terms of duration, in initial stressed words, the duration was almost the same between syllable 1 and 2, but in final stressed words, the duration of the first syllable was shorter than the duration of the second syllable.

2.3 Results

2.3.1 Average F0 ratio

In table 5, the average F0s (absolute value) were provided for the two locations of lexical stress. Nouns with lexical stress on initial and final syllables showed a rise from the initial to the final syllables.

Table 5 Mean Average F0 (Hz) for First and Second Syllables in Minimal Pairs

	syllable 1	syllable 2	differences
initial stress	171	201	-30
final stress	177	211	-34
differences	-6	-10	

In Figure 4, the average F0 (ratio across subjects) is provided for the two different stressed minimal pairs (initial versus final). Repeated Measure ANOVA by subject and item analyses was conducted. The main effect of initial versus final stressed words was not significant in both subject *(F_1 (1, 4) = 0.21, p = 0.67)* and item *(F_2 (1, 5) = 0.39, p = 0. 56)* analyses. The average ratio of F0 (2.95) on the final stressed words was not significantly higher than the average ratio of F0 (2.76) on the initial stressed words. F0 ratios did not differentiate the stressed syllables from the unstressed syllables.

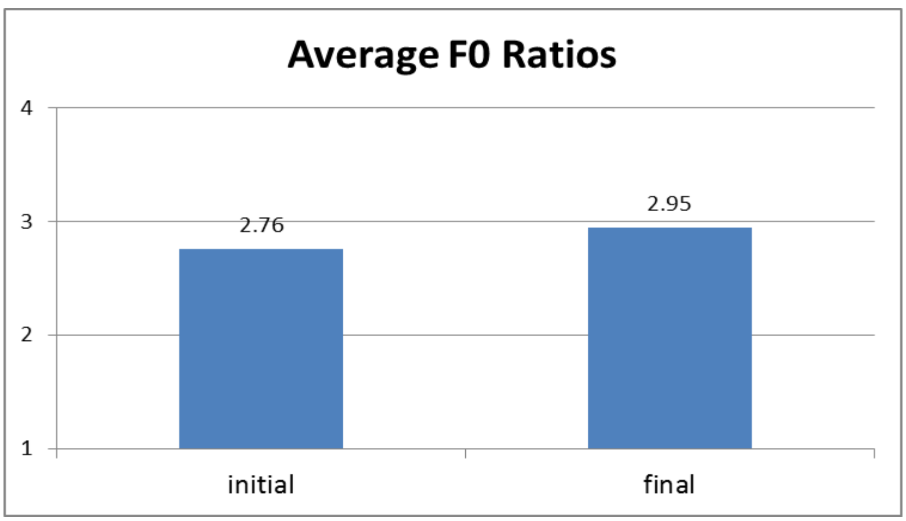

Figure 4 Average F0 ratios initial and final stress stimuli for six minimal pairs.

2.3.2 Duration ratio

In Table 6, the durations (absolute value) were provided for each of the two locations of lexical stress. Nouns with lexical stress on the final and the initial syllables showed slightly different patterns in terms of duration.

Table 6 Mean Duration (ms) for First and Second Syllables in Minimal Pairs

	syllable 1	syllable 2	differences
initial stress	129	126	3
final stress	89	111	22
differences	40	15	

In Figure 5, the durations (ratios across subjects) are provided for each of the minimal pairs. The main effect of the initial stressed versus the final stressed words was significant by the item analysis *(F_2 (1, 5) = 9.43, p = 0. 028)*, and it was marginal significant by the subject analysis *(F_1 (1, 4) = 5.34, p = 0.082)*. The final stressed words have a longer duration in the second syllables than the first syllables, and initial stressed words had a slightly shorter duration on the second syllables than the first syllables. Overall, a significant ratio difference showed that duration does matter in terms of stress location.

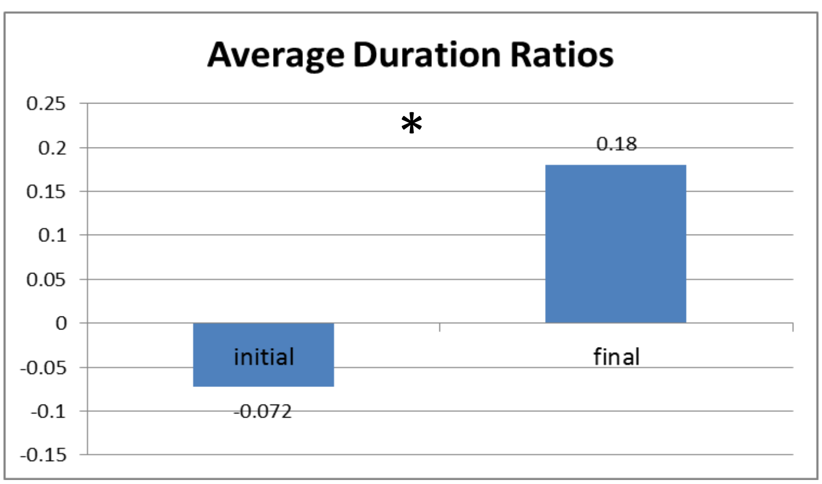

Figure 5 Duration ratios initial and final stress stimuli for six minimal pairs.

2.3.3 Intensity ratio

In table 7, the average intensities were provided for the two locations of lexical stress. Nouns with lexical stress on the final stressed syllables had a slight higher intensity, while nouns with lexical stress on initial stressed words had a slightly lower intensity.

Table 7 Mean the Average Intensity(dB) for First and Second Syllables in Minimal Pairs

	syllable 1	syllable 2	differences
initial stress	69	68	1
final stress	68	69	-1
differences	1	-1	

In Figure 6, the average intensities are provided for the two types of minimal pairs (initial versus final). The main effect of the initial stressed versus final stressed words was marginally significant by the subject analysis *(F_1 (1, 4) = 6.22, p = 0.067)* and significant by the item analysis *(F_2 (1, 5) = 35.27, p = 0. 002)*. This suggested intensity is a strong cue for distinguishing the stressed syllables from the unstressed syllables.

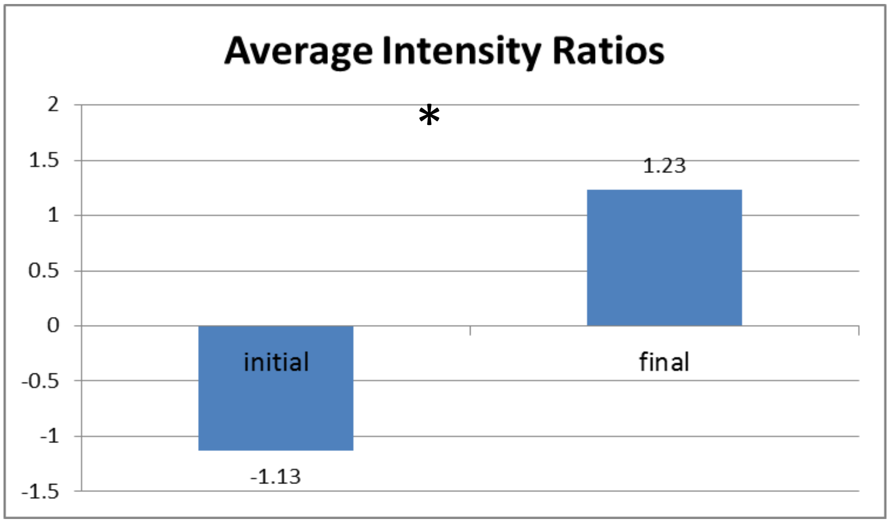

Figure 6 Average intensity ratios initial and final stress stimuli for six minimal pairs

2.4 Discussion

This study investigated the stress pattern in Uyghur using an acoustic analysis of stressed and unstressed syllables in terms of F0, duration and intensity. Acoustic analyses were conducted. In the acoustic analyses, ratios of the parameters were taken into consideration. The results indicated that duration and intensity, but not F0, were the critical cues for assigning stress location. Therefore, Uyghur seems to pattern as a stress-accent language.

In this research, five native Uyghur speakers produced six pairs which contrast in terms of stress placement that included six initial stressed and six final stressed minimal pair words. We examined three acoustic parameters: average fundamental frequency, duration and average intensity. For the analyses, we used ratios of these values as Beckman (1986) did. All stimuli produced by native Uyghur speakers where used and analyzed. The results from the acoustic analyses showed a consistent pattern in which stress was cued by duration and intensity, but not by F0, in Uyghur. We will discuss the results in terms of fundamental frequency, duration and intensity for the ratio analysis data.

For F0, a lack of significant F0 ratios comparing initial stressed words and final stressed words indicated that F0 is not a cue for differentiating the stressed syllables from the unstressed syllables in Uyghur. From the ratio analysis, it is clear that F0 is not a cue for stress location in Uyghur. The rise in absolute F0 values in second syllables regardless of stress type indicated that F0 was not a critical cue for assigning stress location. In other words, speakers raise their F0 in the final syllables in their speech regardless of the stress location. F0 did not distinguish the initial stressed words from the final stressed words.

The second parameter was the duration of each vowel in the stressed and the unstressed syllables in both initial and final stressed words. Duration ratios were compared for the initial stressed and the final stressed words. The duration ratios are significant or marginally significant in both subject and item analyses. Duration ratio was significantly different in initial stressed words compared to final stressed words in the acoustic analysis. These results suggested that duration can distinguish the stressed syllables from the unstressed syllables.

Intensity was the third parameter used to examine stress location in minimal pairs. The intensity ratio was lower in the initial stressed words and higher in the final stressed words. The stressed syllables had a higher intensity than the unstressed syllables both in initial and final stressed words. These results revealed that intensity is a critical cue for differentiating stressed syllables from unstressed syllables.

In sum, in Uyghur, based on ratio analysis for a set of minimal pairs, we concluded that duration and intensity provide cues for stress location. However, fundamental frequency did not differentiate the stressed syllables from the unstressed syllables. Uyghur appear to be a stress-accent language because F0 did not distinguish the stressed syllables from unstressed syllables at all. Because F0 is not used as a stress cue, Uyghur cannot be classified as a pitch-accent language. Moreover, Uyghur uses duration and intensity as cues for distinguishing stressed syllables from unstressed syllables. Therefore, Uyghur appear to pattern consistently as a stress-accent language.

Chapter 3 Experiment 2: An Acoustic Study of Disyllabic Nouns Contrasting in Stress

3.1 Motivation and research questions

3.1.1 Motivation

In Uyghur, minimal pairs contrasting in stem alone are rare. In Experiment 1, one member of the minimal pairs that were used is often borrowed from Arabic or Farsi. For example, in the word, *bala,* 'child,' stress is on the first syllable; however, in *bala,* 'disaster,' which was borrowed from Arabic, the stress is on the second syllable. If both members of the minimal pair in Experiment 1 originated from Turkic words, some letters or syllables were dropped historically and this resulted in a minimal pair. For example, in *Acha* [aʧa[2]], 'elder sister,' the original Turkic version was *aghicha* [aʁiʧa], which means 'a woman servant in a noble or rich house.'(Tahur, personal communication, 2012 March). This older form is no longer used. In current usage, it is shortened to *Acha* [aʧa] with the meaning of 'elder sister'. Other examples included *CHAtaq* [ʧataq], 'bad branch of tree,' which is a word used in our experiment, was originated from *chartaq* [ʧartaq], in which the coda position [r] was dropped. One member of each of five pairs among the six minimal pairs used in Experiment 1 lost some syllables historically, and coincidently became minimal pair counterparts with other members. Finally, some members of the six minimal pairs are much less frequently used, and are only used in dialectical regions or historical usage. For example, *TÖshük* [tøʃyk], as 'kitchen,' is an older version, and it is currently used only in some regions. Therefore, it is possible that some of the minimal pairs used in Experiment 1 are not completely representative of typical lexical items in Uyghur.

---

[2] *aʧa* had the explanation as *eçe:* in Old Turkic in which it has meaning both of sister of mother's and elder sister. (Clauson, 1972, p 20)

Instead of using minimal pairs, the second experiment examined disyllabic words which share the same syllable in the initial position. In these pairs, the first syllable in one member of the disyllabic words is stressed, and in the other member of the disyllabic pairs, the first syllable is not stressed (the second syllable is stressed). For example, in *DAka* 'gauze' versus *daLA* 'plain,' the first syllable 'da' of the first disyllabic word (DAka) is stressed, and the first syllable 'da' in the second disyllabic word (daLA) is not stressed. We can compare the acoustic values in both stressed and unstressed first syllables in those disyllabic words.

One advantage of using these words in Experiment 2 is to examine the influence of syllable structure in stressed syllables. Little research has been done on whether syllable weight matters in Uyghur. Since it is the case that many nouns were borrowed from Arabic and Farsi, and Arabic is a quantity- sensitive stress type language (Altmann, 2006), syllable weight may have an effect. In other words, syllable weight might play a role in assigning stress in Uyghur. In order to control the possible effect of syllable weight in this experiment, we additionally examined two distinct syllabic structures, CVCV versus CVCCVC syllables.

Saimaiti and Feng (2008) categorized syllable structures in Uyghur as V, CV, VC, CVC, CVCC, CCV, CCVCC, CVV, and CVVC based on their corpus study. However, the most common syllable structures are the CV (38.4%) and the CVC (50.33%) in the stem form type and the CV (50.01%) and the CVC (40.44%) in inflectional type, according to Saimaiti and Feng (2008). They used a syllabication algorithm for two different corpuses: one is *Uyghur Tilining Izahliq Lughiti* (Uyghur Explanatory Dictionary, edited by Yakup and Gheyurani, 1999), which has 30,169 words in lemma (stem type), and a second corpus established by Xinjiang University from 2003 to 2006, which has 2,558,810 words. The two corpora include academic articles (15%), newspaper reports and opinions (27%), corporate websites (25%), magazine articles

(13%), and novels (20%). Based on their study, we chose the most common syllable structures (the CV and the CVC) in this research. We assume that if syllable weight has a role in stress assignment, we may have a different pattern across the two groups of words (CVCV words versus CVCCVC words). While Uyghur does have complex codas and onsets, the majority of these is rare and often borrowed from Russian, Arabic, Farsi and Chinese. Saimaiti and Feng (2008) indicated complex codas and onsets are less frequent in Uyghur. The majority of native Turkic words have simple onsets and simple codas.

Experiment 2 also provided the chance to examine the stress pattern on the new set of stimuli. In Experiment 1, we found that duration and intensity, but not F0, provided the cues for stress location. In Experiment 2, we used a new set of stimuli in which the first syllables are contrasted in terms of stress in order to investigate whether we find a similar set of cues to signal stress as we found in Experiment 1.

3.1.2 Research questions and hypotheses

This experiment investigated the following research questions:

(1) Do the acoustic parameters of fundamental frequency, duration, and intensity provide cues for differentiating stressed syllables from unstressed syllables in Uyghur?

Based on the research question, the following hypotheses were tested:

If pitch does matter for stress location, the stressed syllables will have higher pitch than the unstressed syllables. Stressed syllables will be longer than the unstressed syllables if duration is a stress cue. For intensity, if intensity provides the cue for stress location, stressed syllables will have greater intensity than the unstressed syllables.

If the results in Experiment 1 generalize to a new set of stimuli in Experiment 2, duration and intensity will provide the cues for stress assignment in Uyghur. In addition, F0 will not have a role in distinguishing the stressed syllables from unstressed syllables.

To examine syllable structure, we use the CV and the CVC syllable structures. For syllable structure we asked the following research question:

(2) Do distinct syllable structures result in the use of different cues to stress?

If syllable structure matters in Uyghur, the CV and the CVC stimuli may have a different pattern of results. The CV structure may have stronger effects in terms of duration and intensity compared to the CVC structure since the open syllables have longer durations than the closed syllables.

When people are asked to assign stress in Uyghur, there is often inconsistency in assigning stress across different speakers. We also investigated whether the acoustic cues in inconsistently perceived stimuli show fewer stress differences than the acoustically consistently perceived stimuli.

(3) Do the different levels of the perceived consistency affect patterns of stress in Uyghur?

If consistency influences the stress pattern, then consistent stimuli will show systematic use of duration and intensity as cues for stress location, while the inconsistent stimuli will show more variable use of these cues.

3.2 Methods

3.2.1 Stimuli

In the Uyghur dictionary (Yakup and Gheyurani, 1999), stress is not assigned. Given that the stress patterns are not listed, we ran a pilot test in which we examined the perceived stress assignment by native speakers in disyllabic nouns.

In order to select stimuli for Experiment 2, we collected 75 paired words that included 32 CVCV disyllabic word pairs, and 43 CVCCVC disyllabic word pairs. All words are from the Uyghur explanatory dictionary (Yakup and Gheyurani, 1999). When we selected words, first, we tried to choose the disyllabic words that contrasted in the first syllable in terms of stress, but had similar segmental content. Secondly, we avoided the ‘m, n, l, h, r, ng’ sounds following the first syllable target vowels as much as possible, because these sounds make it more difficult to isolate or segment the first syllable.

We used both words in each the pair (150 words), and then randomized the 150 words. We chose 18 native Uyghur participants, who did not participate in any of the production experiments, to evaluate the stress location of the stimuli. The 18 participants were senior-level psychology students from Xinjiang Normal University. We gave the instruction as follows: “please read each word in two different ways: first, put the stress on the first syllable; second, put the stress on the second syllable, and then for each word, make a judgment as to which way is more natural. Choose the one that sounds more natural in terms of the stress assignment. For example, for *paqa* ‘frog,’ read first as *PAqa*; second, read as *paQA*. Decide which one is more natural for you, *PAqa* or *paQA*, in natural speech. If you think *PAqa* sounds more natural than *paQA*, please put 1 (the first syllable is stressed); if you think *paQA* is more natural than *PAqa*, please put 2 (the second syllable is stressed).”

We collected the data from 18 participants for each word since the data showed that stress assignments are not unanimous across for all participants. Therefore, we used a criterion for stress location. If more than 10 among the 18 participants assigned stress on the first syllable in a word, then we chose that word as a first syllable-stressed word. At the same time, we also looked at the counterpart of the selected word that had its first syllable stressed by the majority (10/18) of participants. If the counterpart was assigned as a second syllable-stressed word (meaning that participants put the stress on the second syllable), and left the first syllable as unstressed by more than 10 among the 18 participants, we selected the pair as a target word pair in which two words share the same initial syllables and contrasted in stress location on the first syllables. For example, if the number of the participants who chose *DAka,* 'gauze,' as the first syllable-stressed word was more than 10 among 18, then we chose *DAka* as a first syllable-stressed word. We then looked at its counterpart, *daLA,* 'plain,' whether it was second syllable-stressed or not. If the number of the participants, who chose the counterpart as the second syllable -stressed words, was also more than 10 among 18, then we chose both members of the pair as target words.

Based on these criteria, we chose six word pairs with CVCV structure (as shown in Table 8 below). The CVCV structure (I will refer those stimuli as the CV-consistent in the following sections) includes disyllabic words in which both of the first and second syllables have a CV structure. Among the six word pairs, each pair included the first syllable-stressed word versus the first syllable-unstressed word. Participants who identified the first syllable-stressed on the first member also consistently identified the second syllable-stressed on the second member of the pair.

For the CV-consistent stimuli, we found that participants' ratings of the consistency on the stress locations between the first syllables and the second syllables were not significant *(t (5) = 1.965, p = 0.107),* which indicated that participants' stress assignment for the stimuli with the first syllable-stressed (the mean was 14.33) was similar to the stimuli with the second syllable-stressed (the mean was 12.83). Both members of the target pair showed consistent assignment of stress location.

Table 8 The CV consistent Stimuli (participants' assignment of stress)

first syllable stressed			second syllable stressed		
word	ratings	English gloss	word	ratings	English gloss
DAka	14/18	gauze	daLA	15/18	plain
CHAsa	14/18	square	chaNA	14/18	sledge
BAza	16/18	base	baHA	13/18	price
DAcha	14/18	villa	daDA	13/18	father
DOra	14/18	medicine	doQA	12/18	forehead
POchi	14/18	boaster	poTA	10/18	waistbelt

For the CVCCVC, we chose six word pairs in with CVCCVC structure (as shown in Table 9 below). The same method for choosing pairs in the CV-consistent group was used in the CVCCVC structure. We matched the first syllable-stressed of one member of the pair with the second syllable-stressed second member of the pair.

For the CVC-consistent stimuli, we found that the ratings of consistency between the first syllables and the second syllables were not significant *(t (5) = 1.12, p = 0.314)*, which indicated that participants' stress assignment for the stimuli with the first syllable-stressed (the mean was 11.83) was similar to the stimuli with second syllable-stress (the mean was 10.83). In other words, the first syllable-stressed words and the second syllable-stressed words were consistent in terms of stress assignment.

Table 9 The CVC-consistent Stimuli (participants' assignment of stress)

first syllable stressed			second syllable stressed		
word	ratings	English gloss	word	ratings	English gloss
SEPkün	11/18	mole	sepDASH	13/18	fellow
PUTbol	11/18	football	putLASH	12/18	trapping
KASsir	12/18	accountant	kasTUM	10/18	suit
PARnik	11/18	greenhouse	parLASH	10/18	choosing
BANkir	14/18	banker	banTIK	10/18	tie
TOQmaq	12/18	stick	toqQUZ	10/18	nine

We separately compared the ratings of stress assignment to the first syllables and the second syllables that were stressed between the CV-consistent group and the CVC-consistent group. We found that the mean difference of first syllables on the CV-consistent group and the CVC-consistent group was significant *($t(5) = 4.44$, $p = 0.007$)*. The mean difference on the second syllables was also significant *($t(5) = 1.47$, $p = 0.007$)*. The ratings of the CV structure were significantly different from the ratings of the CVC structure. This indicated that the CV-consistent structure stimuli were less variable when participants were assigning stress location compared to the CVC-consistent structure stimuli.

A final set of stimuli were chosen investigating whether the inconsistency in stress assignment affect the acoustic parameters speakers use when assigning stress. We chose 6 word pairs with the CVCV structure for the CV-inconsistent group. For these stimuli, fewer than 9 of 18 participants assigned stress on the first syllables, but more than 10 of 18 participants assigned the stress on the second syllables in the counterparts (as shown in Table 10 below). In other words, the first syllable-stressed words are not consistently perceived as first syllable stressed ones across speakers. Less than 50% of speakers indicated the stimuli with first syllable-stress.

Table 10 The CV inconsistent Stimuli (participants' assignment of stress)

First syllable stressed			second syllable stressed		
word	ratings	English gloss	word	ratings	English gloss
SOda	9/18	business	soQA	13/18	plough
KIno	9/18	movie	kiSHI	11/18	person
JAza	8/18	punishment	jaPA	15/18	hardness
SHIre	8/18	desk	shiPA	14/18	cure
CHOla	7/18	free time	choKA	12/18	chopstick
TOxu	8/18	hen	toPA	12/18	soil

For the CV-inconsistent stimuli, we found that the ratings on the consistency between the first syllables and the second syllables assignment were significantly different *($t(5) = -6.53$, $p = 0.01$)*, which indicated that participants' stress assignment for the stimuli with the first syllable-stressed (the mean was 8.12) was different from the stimuli with the second syllable-stressed (the mean was 12.83). In other words, more participants who assigned stress on the second syllables in the second syllable-stressed group are more consistent than the ones in first syllable-stressed group. Ratings for CV-inconsistent stimuli were lower for the first syllable than for the second syllable stressed items.

We also compared the stress assignment for the first syllables in both the CV-consistent and the CV-inconsistent groups (Table 8 and Table 10), and found that participants' ratings were significantly different for the first syllable in the CV-consistent group compared to the CV-inconsistent group *($t(5) = 12.92$, $p = 0.0001$)*. The significant difference in the consistent versus the inconsistent groups indicated that participants assigned stress differently on the first syllables in the CV-consistent group and the CV-inconsistent group.

In summary, when choosing the stimuli, we contrasted the first syllable-stressed versus first syllable unstressed words (second syllable-stressed words) in disyllabic words with segmentally identical first syllables. We also compared stimuli with different syllabic structures

(CV-consistent versus CVC-consistent) and examined stimuli with consistent versus inconsistent perceptual stress assignment (CV-consistent versus CV-inconsistent).

3.2.2 Participants

Seven male native Uyghur speakers living in the USA were recorded individually. The range was from 24-35 years old. See the Table 11 for information about the participants. They learned English as third language at the age 16 (AOA: 13-18 years). All of them are graduate students at the universities in the USA. All are majoring in Engineering or Science.

Table 11 The Background Information of Participants

Participants	English level	years in USA	AOA	Chinese	Other languages
1	near-native	6	18	near-native	
2	near-native	3.5	13	near-native	near-native Turkish
3	good	0.6	15	near-native	near-native Uzbek
4	good	1.5	16	near-native	near-native Uzbek
5	good	1.5	15	near-native	
6	good	2.5	15	good	
7	near-native	4.5	15	near-native	

AOA: Age of acquisition; English and Chinese levels were evaluated by participants themselves.

Each participant's speech sounds were recorded with a Marantz PMD 671 solid-state recorder via an Electro-Voice ND767a microphone.

3.2.3 Procedure

Each word was recorded in a fixed sentence condition. Stimuli are recorded in a condition in which target words are set in a fixed sentence: "Men hazir ____ deymen."(I will say ____ now.). The fixed condition was used for data analysis. The following examples of target words 'DAka' (example 3) and 'daLA' (example 4) are shown in the fixed condition:

(3) the fixed sentence for first syllable-stressed words

Men	hazir	***DAka***	de-y-men.
I	now	gauze	say-PRE-1^{st} sg

'I will say gauze now.'

(4) the fixed sentence for first syllable unstressed swords

Men	hazir	***daLA***	de-y-men.
I	now	plain	say-PRE-1^{st} sg

'I will say plain now.'

Each fixed sentence containing the target stimuli was repeated three times in different random orders across all three groups of stimuli. A total of 108 tokens were analyzed for each speaker. The total tokens are 756 (7 speakers × 18 pairs × 2 words × 3 repetitions).

3.2.4 Measurements

In this experiment, we measured mean fundamental frequency, duration and mean intensity based on the vowel, similar to Experiment 1. We compared the first syllables in the members of disyllabic words as shown in the example (3) and (4) above 'DAka' versus 'daLA'. Instead of using ratio analysis, we performed a direct comparison of the first syllable using absolute values. In other words, we compared three different parameters (mean F0, duration, and mean intensity) of the first syllable ($\sigma 1$) in 'DAka' with the first syllable ($\sigma 1$) of 'daLA.'

3.2.5 Statistics

We used Repeated Measure ANOVA by subjects and by items. In the subject analysis, we averaged separately all items' values in the CV-consistent, the CVC-consistent and the CV-inconsistent groups. The subject analysis showed the subjects' behaviors across items in the stressed and unstressed syllables. In the item analysis, we averaged all subjects on each item.

Item analysis indicated how each item's behavior changes across subjects in the different conditions.

First, we conducted a Repeated Measure (3x2) ANOVA, in which the group has three levels (CV-consistent, CVC-consistent, and CV-inconsistent) and the stress condition has two levels (first syllable-stressed and first syllable-unstressed) for all subjects and items separately. For the Repeated Measure Analysis, we treated groups (syllable structure or consistency) and stress condition as within subject variables for the subject analysis; however, for the item analysis, we treated stress condition as a within subject variable, and groups (syllable structure or consistency) as between subject variables, respectively.

Secondly, we conducted a CV-consistent versus a CVC-consistent repeated Measure (2x2) ANOVA in which syllable structure has two levels (the CV-consistent and the CVC-consistent), and stress condition has two levels (the stressed syllables and unstressed syllables). This analysis evaluates whether the syllable structure affects the stress patterns by comparing the CV-consistent and the CVC-consistent stimuli. In other words, we manipulated the syllable structure. Differences would be due to difference in syllable structure.

In order to investigate the consistency effect of perceiving stress location, we used Repeated Measure (2x2) ANOVAs in which the consistency has two levels (CV-consistent versus CV-inconsistent) and stress has two levels (stressed syllables versus unstressed syllables). The purpose of doing this analysis is to examine whether consistency affects the stress pattern in stimuli with a CV syllable structure. In other words, we controlled the syllable structure and manipulated the perceived consistency.

3.3 Results

We measured the mean F0, duration, and the mean intensity of the vowels in the first syllables of the disyllabic words in stressed and unstressed conditions. We removed two pairs in the CV-inconsistent group and one pair in the CVC group because high vowels occurring between voiceless consonants or preceding voiceless consonants were often devoiced. For the first CV-inconsistent case, the word *kiSHI* [kiʃi], 'person,' has stress on the second syllable, *SHI*, and the first syllable, *ki*, was unstressed. The /i/ after voiceless consonants [k] was devoiced. If one member of a pair was eliminated from the data, we also removed the counterpart (*kino* 'movie', in which the *KI* was stressed and *no* was unstressed). The other case in the CV-inconsistent group was *SHIre* [ʃiɹɛ], 'desk,' and the counterpart was *shiPA* [ʃipa], 'cure.' In both of them, the first vowel /i/ was often devoiced. The last case was from the CVC-consistent group in which *PUTbol* [pʰutbol], 'football,' and its counterpart *putLASH* [pʰutlaʃ] 'trapping,' were eliminated from the data analysis. The devoicing of [u] occurred for all subjects. Therefore, we analyzed six word pairs in the CV-consistent group, five word pairs in the CVC-consistent group and four word pairs in the CV- inconsistent group.

We had three groups: the CV-consistent, the CVC-consistent and the CV- inconsistent groups. We show the results based on the F0, duration and intensity separately. First, we present overall results, then the syllable structure analysis and finally the consistency results.

3.3.1 Average F0

3.3.1.1 Overall results

As shown in Table 12 below, the average F0 (absolute value) was provided for the first syllables of the word pairs in three groups. The main effect of stress was not significant in either subject analysis *(F_1 (1, 6) = 0.74, p = 0.422)* or item analysis *(F_2 (1, 12) = 1.44, p = 0.253)*. No significant difference was found between the stressed syllables and the unstressed syllables regardless of the group. The results showed that F0 cannot differentiate the stressed syllables from the unstressed syllables.

The main effect of the groups was significant by the subject analysis *(F_1 (2, 12) = 15.67, p = 0.0001)* and marginally significant by the item analysis *(F_2 (2, 12) = 3.53, p = 0.062)*. This indicated that the overall F0 patterns for the three groups were different. In other words, the groups (*121 Hz* for CV-consistent, *126 Hz* for CVC-consistent and *123 Hz* for CV-inconsistent) were different from each other with the CVC consistent stimuli showing higher F0. The interaction between the groups and stress types was not significant either by the subject analysis *(F_1 (2, 12) = 2.81, p = 0.100)* or by the item analysis *(F_2 (2, 12) = 1.26, p = 0.318)*.

Table 12 F0 (Hz)Values and Standard Deviations among the Groups

Stress type	CV-con	CVC-con	CV-incon	overall AVG
stressed	121 (17.83)	125 (19.48)	122 (19.35)	123
unstressed	121 (17.74)	126 (20.11)	123 (19.46)	123
overall AVG	121	126	123	

Number in parentheses indicate standard deviation

From the overall analysis of F0, we found no significant differences across the stressed syllables and the unstressed syllables.

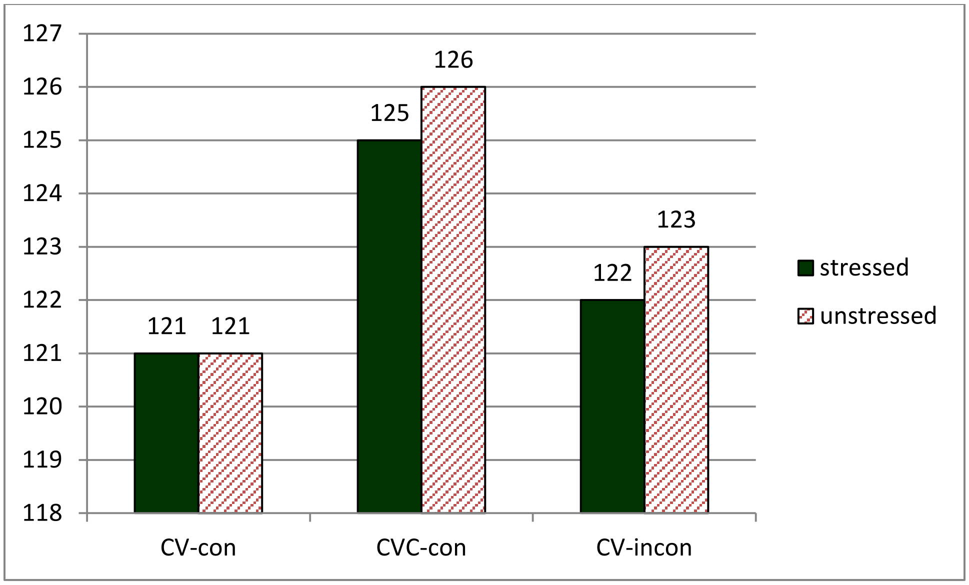

Figure 7 Average F0 across three types of stimuli: dark bar shows the data for stimuli with the first syllable stressed and the crossed bar shows the data for the stimuli where the first syllable of the counterpart is not stressed. X axis shows three groups and Y axis is F0 values in Hz.

From the F0 analysis, we concluded that F0 cannot distinguish the stressed syllables from unstressed syllables across the three groups of stimuli. The overall analysis showed that the main effect of stress was not significant. Therefore, we can claim that F0 is not a cue to indicate stress location in Uyghur.

3.3.2 Duration

(1) Overall results

As shown in Table 13 below, the average duration (absolute value) was shown for the first syllables of the word pairs in the three groups of stimuli. The main effect of stress was significant in both the subject analysis (F_1 *(1, 6) = 38.96, p = 0.001)* and the item analysis *(*F_2 *(1, 12) = 44.80, p = 0.0001)*. The stressed syllables *(106ms)* had longer duration than the unstressed syllables *(76ms)* regardless of the groups. From the significant stress type effect, we can conclude that duration can differentiate the stressed syllables from the unstressed syllables.

The main effect of the group was also strongly significant in both the subject analysis *(F_1 (2, 12) = 50.19, p = 0.0001)* and the item analysis *(F_2 (2, 12) = 13.37, p = 0.001)*. A significant main effect of groups indicated that the patterns of the three groups of stimuli were different regardless of stress type. In other words, groups of stimuli (*112ms* for the CV-consistent, *75ms* for the CVC-consistent and *86ms* for the CV-inconsistent) were different from each other in which the CV-consistent group had longer duration than CVC-consistent group *($p < 0.001$)* and CV-inconsistent group *($p = 0.006$)*, while CVC-consistent group was not different from CV-inconsistent group *($p = 0.204$)*. From the overall analysis of duration, we discovered that the stress effect and the group effect were significant, which not only indicated that the stressed syllables were significantly longer than the unstressed syllables but also groups were significantly different from each other. The interaction between the group and stress type was also significant in both the subject analysis *(F_1 (2, 12) = 37.08, p = 0.0001)* and the item analysis *(F_2 (2, 12) = 11.07, p = 0.002)*. In terms of duration, the degree of the change from the stressed syllables to the unstressed syllables was different across groups.

Table 13 Duration (ms)Mean Values and Standard Deviation among the Groups

Stress type	CV-con	CVC-con	CV-incon	overall AVG
stressed	138 (25.43)	77 (8.8)	104 (24.74)	106
unstressed	86 (10.96)	73 (6.2)	68 (8.76)	76
overall AVG	112	75	86	

Numbers in parentheses indicate standard deviation

We also ran a paired t-test for the stressed syllables and the unstressed syllables for each group in terms of duration. In the CV-consistent group, a significant difference was found between the stressed syllables *(138ms)* and the unstressed syllables *(86ms)* in both the subject analysis (t_1 *(6) = 7.41, p = 0.0001)* and the item analysis (t_2 *(5) = 6.29, p = 0.001)*. In the CVC-consistent group, a marginal significant difference between the stressed syllables (*77ms*) and the unstressed syllables (*73ms*) was found by the subject analysis *(t_1 (6) = 2.14, p = 0.076)*, but was not significant by the item analysis *(t_2 (4) = 1.41, p = 0.232)*. In the CV-inconsistent group, a significant difference in both the subject analysis *(t_1 (6) = 4.92, p = 0.003)* and the item analysis *(t_2 (3) = 3.27, p = 0.047)* demonstrated that the stressed syllables (*104ms*) had longer duration than the unstressed syllables *(68ms)*. Overall, the CV stimuli showed more systematic differences between the stressed syllables versus the unstressed syllables.

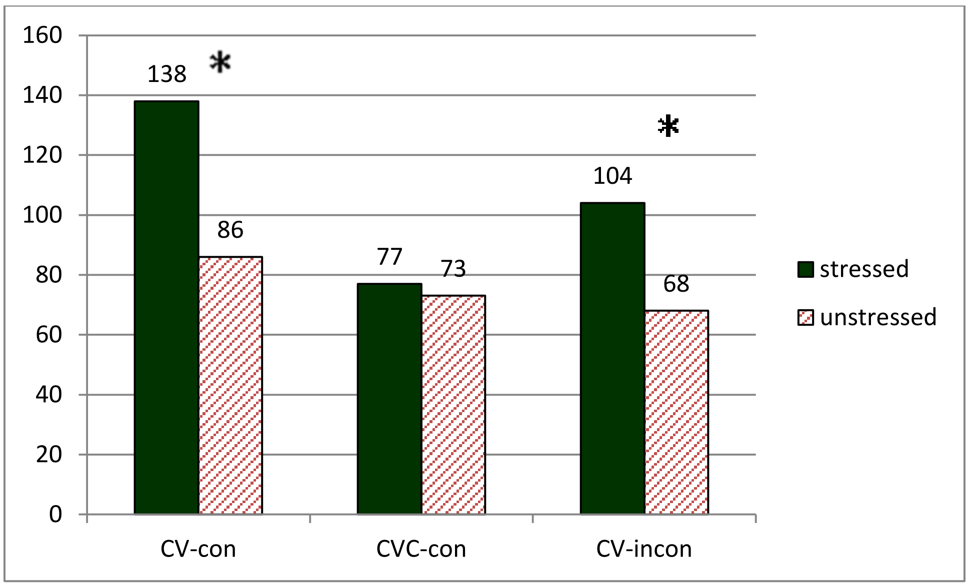

Figure 8 Average duration (ms) across three types of stimuli: dark bar shows the data for stimuli with the first syllable stressed and the crossed bar shows the data for the stimuli where the first syllable of the counterpart is not stressed. X axis shows three groups and Y axis is duration values in ms.

We also conducted separate syllable structure and consistency analyses in order to examine syllable structure and consistency effects.

(2) Syllable structure

In order to examine the syllable structure effect, we grouped the CV-consistent and CVC-consistent groups, and used a two-way Repeated Measure ANOVA. The main effect of stress-type was significant in both the subject analysis *(F_1 (1, 6) = 50.41, p = 0.0001)* and the item analysis *(F_2 (1, 9) = 34.68, p = 0.0001)*. Regardless of syllable types, duration for the stressed syllables (*107ms)* was significantly longer than for the unstressed syllables (*80ms)*. The results suggested that duration does provide information about stress location in disyllabic nouns. Overall, stressed syllables were longer than unstressed syllables.

Moreover, the main effect of syllable type was significant in both the subject analysis *(F_1 (1, 6) = 82.85, p = 0.0001)* and the item analysis *(F_2 (1, 9) = 26.62, p = 0.001)*. The CV syllable structure (*112ms*) had a longer duration than the CVC syllables (*75ms*) regardless of stress type. The result is obvious, because the vowels in open syllables (CV) have a longer duration than the ones in closed syllables (CVC).

The interaction between stress type and syllable type was significant in both the subject analysis *(F_1 (1, 6) = 56.02, p = 0.0001)* and the item analysis *(F_2 (1, 9) = 27.18, p = 0.001)*. The pattern of duration changes from the stressed syllables to the unstressed syllables was different across the CV and the CVC groups. In other words, the degree of duration changes between the stressed syllables and the unstressed syllables was bigger in the CV syllable stimuli than in the CVC syllable stimuli.

From the syllable structure analysis, we concluded that different syllable structures (the CV-consistent versus the CVC-consistent) have different patterns across the stressed syllables and unstressed syllables. Duration changes according to stress type and also according to syllable structure type.

(3) Consistency effects

In order to investigate consistency, we used the CV-consistent and the CV-inconsistent groups, and used a two-way Repeated Measure ANOVA. The main effect of stress-type was significant in both the subject analysis *(F_1 (1, 6) = 40.89, p = 0.001)* and the item analysis *(F_2 (1, 8) = 42.37, p = 0.0001)*. A significant main effect of stress type showed that the stressed syllables *(121 ms)* were significantly longer than the unstressed syllables *(77ms)*. The duration effect on the stressed syllable compared to unstressed syllables further supported the pattern that duration does provide information about stress location in the same syllable structures even when differing on evaluated perceptual consistency of disyllabic nouns. Moreover, the main effect of consistency was significant in both the subject analysis *(F_1 (1, 6) = 50.73, p = 0.0001)* and the item analysis *(F_2 (1, 8) = 10.80, p = 0.011)*. A significant consistency effect showed that the CV-consistent syllables (*112ms*) had a longer duration than the CV-inconsistent (*86ms*) regardless of stress type. However, the interaction between stress type and consistency was significant by the subject analysis *(F_1 (1, 6) = 17.23, p = 0.006)*, but not by the item analysis *(F_2 (1, 8) = 1.48, p = 0.258),* suggesting that duration differences were similar across stimuli varying in consistency.

From the results of the overall analysis, syllable structure and consistency analyses, we concluded that duration can differentiate the stressed syllables from the unstressed syllables. In other words, the stressed syllables had a longer duration than the unstressed syllables, which indicated that duration is a cue to indicate stress location in Uyghur. While duration is a cue, it is not used as systematically in certain syllable structures (the CVC syllable structure). Consistency did not systematically affect use of duration as a cue to stress.

3.3.3 Intensity

The average intensity (absolute value) was provided for the first syllables of the word pairs in the three groups of stimuli. The main effect of stress was not significant in either the subject analysis *(F_1 (1, 6) = 1.44, p = 0.276)* or the item analysis *(F_2 (1, 12) = 3.44, p = 0.088)*. Across the groups, no significant difference between the stressed syllables *(59dB)* and the unstressed syllables *(58dB)* revealed that the role of intensity does not seem to differentiate the stressed syllables from the unstressed syllables. The main effect of groups was marginally significant by the subject analysis *(F_1 (2, 12) = 3.38, p = 0.069)*, but not significant by the item analysis *(F_2 (2, 12) = 2.38, p = 0.135)*. This indicated that the patterns across the three groups of stimuli were not different. Further, the interaction between the group and stress type was not significant in either the subject analysis *(F_1 (2, 12) = 2.80, p = 0.100)* or the item analysis *(F_2 (2, 12) = 1.76, p = 0.214)*.

Table 14 Intensity(dB) Mean Values and Standard Deviation among the Groups

Stress type	CV-con	CVC-con	CV-incon	overall AVG
stressed	59 (4.3)	59 (3.46)	58 (4.15)	59
unstressed	58 (3.81)	59 (3.15)	58 (3.6)	58
overall AVG	58.5	59	58	

Numbers in the parentheses indicate standard deviation

From the overall analysis of intensity, we found that the stress type effect and group effects were not significant suggesting that intensity does not play a role in indicating stress in Uyghur.

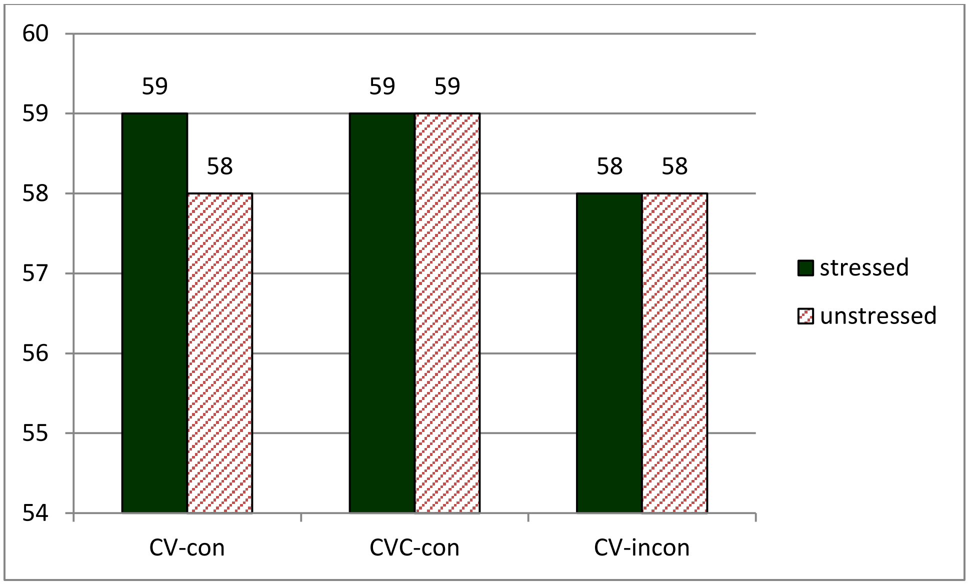

Figure 9 Average intensity (dB) across three types of stimuli: dark bar shows the data for stimuli with the first syllable stressed and the crossed bar shows the data for the stimuli where the first syllable of the counterpart is not stressed. X axis shows three groups and Y axis is intensity values in dB.

The results showed that intensity does not differentiate the stressed syllables from the unstressed syllables. Intensity effects were not consistently observed across the subject analysis and the item analysis. Overall, no intensity effect was found in Experiment 2.

3.4 Discussion

3.4.1 Statement

The purpose of this experiment was to discover how the Uyghur native speakers use acoustic parameters to distinguish the stressed syllables from the unstressed syllables in continuous speech. In Experiment 2, we used disyllabic nouns that contrasted on the first syllables in terms of stress location. Another goal of this experiment was to reexamine the results in Experiment 1 on the new set of stimuli. Since we did not use minimal pairs, we used absolute values of identical first syllables that were either stressed or unstressed. In addition, we manipulated the syllable structure as well as the consistency of stress assignment. Consequently, we had the CV-consistent, the CVC-consistent and the CV-inconsistent groups. Similar to Experiment 1, we measured fundamental frequency, duration and intensity on the vowels in the first syllables. In Experiment 2, we found that duration was a strong cue for stress location in all groups. Intensity and F0 were not cues for distinguishing the stressed syllables from the unstressed syllables. We discuss the results in terms of fundamental frequency, duration and intensity, separately.

The first parameter examined was F0 in stressed and the unstressed first syllables in disyllabic nouns. From the overall analysis, F0 was not different in the stressed and the unstressed syllables across all stimuli groups. In addition, we did not find a syllable structure effect or a consistency effect. The stress patterns were the same across groups (CV-consistent, CVC-consistent and CV-inconsistent). Therefore, it is clear that F0 is not a cue for the stress location in Uyghur.

Duration was examined in the first syllables of the disyllabic words in which the first syllables were contrasted in terms of stress. The duration of the stressed syllables was significantly longer than the unstressed syllables across stimuli groups. When we manipulated the syllable structure, duration was different not only by the stress locations, but also by syllable structures (the CV and the CVC syllable structures). The role of duration was less prominent in the CVC syllable structure. In other words, the CV-consistent stimuli showed greater effects across stressed versus unstressed syllables, while the CVC-consistent group showed less of an effect in duration. The CV stimuli had a stronger duration effect on stressed syllables. In addition, when we were manipulating the different levels of consistency in assigning the stress, duration did not differ. In conclusion, changes in duration were related to stress location, and these effects varied across syllable type. The stressed syllables had longer duration than the unstressed syllables, showing the most sizeable effects for the CV stimuli. Duration is a strong cue for stress location in Uyghur.

Intensity was the third parameter used to test the stress location in disyllabic words that contrasted stress on the first syllables. From the overall results, intensity in the stressed syllables was not significantly different from the unstressed syllables. Moreover, intensity did not change according to syllables structure or consistency. In subject analysis, only the CV-consistent group showed minor effects of intensity on the stress location due to two participants who differentiated the stressed syllables from unstressed syllables in terms of intensity. Overall, intensity does not appear to be a cue for stress location in Uyghur.

3.4.2 Comparison of the results in minimal pairs (Experiment1)

In Experiment 1, we compared the stressed syllables versus the unstressed syllables in minimal pairs. We found that duration and intensity were strong cues for stress location, but not F0. In Experiment 2, in which we used disyllabic word pairs in which stress was contrasted on the first syllables, we also found the duration was a strong cue, but intensity and F0 were not strong cues for stress location. We also examined syllable structure and consistency effects. In Experiment 2, duration was a strong stress cue for the open syllable structure (CV). Stress cues for duration did not differ depending on perceived stress. Unlike Experiment 1, intensity was not a strong cue in Experiment 2. However, similar to Experiment 1, F0 still did not provide any information in terms of stress location in Uyghur.

Chapter 4 Experiment 3 An Interaction Between Lexical Stress and Sentential Intonation

4.1 Motivation and research questions

4.1.1 Motivation

Lexical stress is embedded in sentential intonation. When people perceive the prominence of words or syllables in sentences, sentential intonation may overlay and interface with perception of lexical stress. Acoustically, lexical stress can be observed by changes in fundamental frequency, duration and intensity. Intonation includes speakers' emotion and attitudes. Sentence intonation variations can also be indicated by fundamental frequency, duration and intensity (Beckman, 1986). Given this, cues to lexical stress may change depending not only on the position of the target words in the sentences, but also on overall intonation contour. This chapter will deal with the interaction of sentential intonation with lexical stress. In order to examine sentential intonation in an approachable way, we will contrast declarative and question intonation. To control location of the stressed word in the sentence, declarative assertion and declarative question sentences will be contrasted in which the word order is identical so that the target word is in exactly the same location and same context and only the intonation contour (declarative assertion versus declarative question) varies.

In most languages, a raised pitch at the end of sentence indicates that an utterance is intended as a question compared to a lowed pitch that indicates a declarative sentence (Hirst and Cristo, 1998). F0 contour is strongly correlated to intonation. F0 contour is usually falling if the declarative sentence is an assertion of a fact or event. If the sentence is a question, the F0 contour often rises at the end of the sentence (Hirst and Cristo, 1998). In Experiment 3, we examine the interaction of lexical stress with sentential intonation using declarative sentences as assertions and declarative sentences as questions. Hirst et al, (1998) indicated that F0 might play a greater role and consequently combine with duration and intensity cues for differentiating accented syllables from unaccented syllables at the lexical level, especially in final position (at the end of sentence in which F0 may be falling or rising most prominently). A second possibility is that F0 may not play a role in final position, with the cues for lexical stress similar for both declarative assertion and declarative question sentences.

In order to find additional evidence for the role of F0, Lindström & Remijsen (2005) examined the Kuot language in which they used eight minimal and near-minimal pairs (disyllabic or three syllabic words) that were embedded in four different utterance positions which have four different F0 counters: (1) sentence initially, not followed by an intonationally marked phrase boundary; (2) sentence- medially, not followed by an intonationally marked phrasal boundary; (3) sentence medial, with an intonational phrase marked boundary, and (4) sentence finally. They measured average F0, duration, F1 and F2 for each member of minimal pairs. In all cases, no interaction existed between intonational contour and duration and vowel quality. Moreover, there was no F0 role for distinguishing stressed syllables from unstressed syllables. In other words, there was no interaction between duration and vowel quality with sentence position. This implied that even though there was a significant F0 change, the cues to lexical stress remained the same. Thus, F0 at the sentential intonation was separated from F0 at the lexical level.

In Experiment 1and Experiment 2, we showed that F0 did not provide a cue for stress in Uyghur. In Experiment 1, duration and intensity were stress cues, and in Experiment 2, duration was a stress cue. Uyghur differs from English in which duration, intensity and F0 all provide cues for stress in English. Uyghur is also different from Japanese in that only F0 provides a cue for stress location in Japanese. Uyghur might share some characteristics with Kuot, which uses duration and vowel quality, but not F0. This chapter will provide a window for testing interaction of the sentential and lexical level intonation cues. In addition, by providing distinct F0 contours using contrasting sentential contours, the interaction between lexical F0 and sentential F0 was examined.

4.1.2 Research questions and hypothesis

The research question is as follow:

Do the acoustics parameters indicating lexical stress change under the different sentential intonation patterns?

Hypothesis 1: If the declarative assertion type and declarative question type sentence are different in terms of F0, there will be an overall effect of sentence type (declarative assertion and declarative question) in terms of F0.

Hypothesis 2: If there is an interaction between sentential intonations (declarative question and declarative assertion sentences) with lexical stress, we will find effects for duration, intensity and F0 for distinguishing the stressed syllables from the unstressed syllables across sentence type. Two possibilities may occur: (1) an enhancement effect that indicates these parameters (duration, intensity and F0) will be heightened by sentential intonation; (2) a diminished effect that indicates that these parameters (duration, intensity and F0) will be minimized by the changes in sentential intonation.

4.2 Methods

4.2.1 Stimuli

The stimuli were the same the CV-consistent disyllabic nouns used in Experiment 2 (see Table 8) including six word pairs with first syllables stress and second syllable stress. We selected two sentences: one was a declarative sentence with assertion; another was a declarative sentence with question. In English, the sentence 'Jon bought a book.' is a declarative sentence with assertion; on the other hand, the sentence as 'Jon bought a book?' is a declarative sentence with question. In Uyghur, declarative assertion and declarative question sentence are similar to English. We will use these two types of sentences and contrast disyllabic nouns in final sentence position.

4.2.2 Participants

Six native (male) Uyghur speakers living in the USA were recorded individually similar to Experiment 2. Subjects were instructed to read the sentences in each condition (declarative assertion and declarative question) for each member of the pair in a random order three times. Each participant's speech sounds were recorded with a Marantz PMD 671 solid-state recorder. The speakers were recorded in a quiet room with an Electro-Voice ND767a microphone. The recordings were digitized at a sampling rate of 22.05 kHz.

4.2.3 Procedure

Each word is recorded in two different sentential conditions. The followings are examples of the stimuli in Experiment 3: examples of a declarative assertion (5a) and a declarative question (5b) for first syllable-stressed disyllabic nouns; examples of a declarative assertion (6a) and a declarative question (6b) for first syllable unstressed disyllabic nouns.

Declarative assertion (first syllable-stressed)

(5a) Biz-ge kérek bol-ghan-i chong we pakiz DAka
we-DAT necessary be-ADJL-ACC[3] big and clean gauze
'What we need is big and clean gauze.'

Declarative question (first syllable-stressed)

(5b) Biz-ge kérek bol-ghan-i chong we pakiz DAka
we-DAT necessary be-ADJL-ACC big and clean gauze
'What we need is big and clean gauze?'

Declarative assertion (first syllable-unstressed)

(6a) U-ning yil-lar-che siz-ghan resim-lar-i güzel bir daLA
He/She-GENyear-PL-LIM draw-ADJL painting-PL-POSS beautiful one countryside
'The paintings she draws every year are of a beautiful countryside.'

Declarative question (first syllable-unstressed)

(6b) U-ning yil-lar-che siz-ghan resim-lar-i güzel bir daLA
He/She-GENyear-PL- draw-ADJL painting-PL-POSS beautiful one countryside
'The paintings she draws every year are of a beautiful countryside.'

All target words were put into the two sentential conditions that were declarative assertion and declarative question types. Sentences were randomized. The CV-consistent group stimuli were used. Therefore, for the six word pair stimuli, we had 72 sentence tokens (6 pairs × 2 words × 2 sentence conditions × 3 repetitions) for each participant. Three repetitions of the stimuli were recorded. The total will be 432 (6 speakers × 6 pairs × 2 words × 2 sentence conditions × 3 repetitions) tokens.

4.2.4 Measurements

In this data set, we measured averaged fundamental frequency, duration and intensity of the first syllables in disyllabic words, identical to the measurements used in Experiment 2. We compared the first syllable (da) in disyllabic words as shown above ('DAka' versus 'daLA') in both

[3] ADJL represents adjectilizer, the suffix GAn converts the verb as adjective-like form; LIM represents limitative case. DAT represents dative case; ACC represents accusative case; GEN represents genitive case; PL represents plural in nouns; POSS represents possessive case

declarative assertion and declarative question types of sentences. Instead of using ratio analysis, we adopted the direct comparison with absolute values. In other words, we compared three different parameters (F0, duration and intensity) for the first syllables of the words 'DAka' and 'daLA'.

4.2.5 Statistics

For statistics, we used Repeated Measure ANOVAs similar to Experiment 2. In Experiment 3, we also had a within subject variable that was sentence type which had two levels as declarative assertion and declarative question types. We also used subject and item analyses. Analyses included overall analysis, syllable structure effects analysis and consistency analysis, similar to Experiment 2.

We used the two way Repeated Measure ANOVA (2(sentence type) x 2(stress type)) in which the sentence types (declarative assertion versus declarative question), and stress types (first syllable stressed versus first syllable unstressed) were within subject variables for subject analysis and for item analysis.

4.3 Results

For Experiment 3, the target word was in final position for the two sentences types: declarative assertion and declarative question types. The results were based on the average F0, duration and the average intensity shown below.

4.3.1 Average F0

F0 was measured in the first syllables of disyllabic nouns that contrasted on the first syllables in terms of stress. The data for each sentence type and stress type was shown in Figure 10.

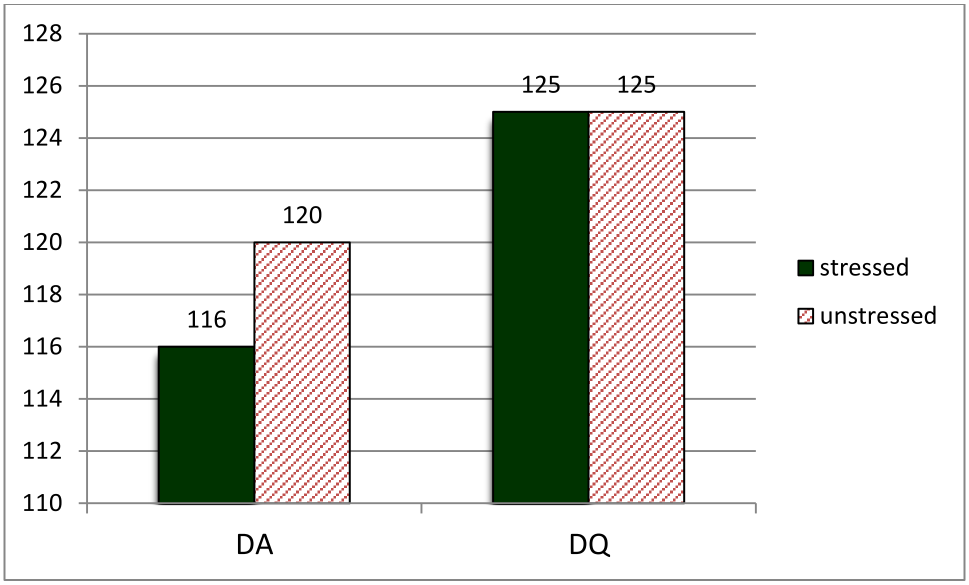

Figure 10 Average F0 (Hz) for the CV-consistent stimuli group: Dark bar shows data for the stimuli with first syllable stressed and the crossed bar shows data for the stimuli where the first syllable of the counterpart is not stressed. X axis shows declarative assertion (DA) versus declarative question (DQ) condition and Y axis is F0 values in Hz.

No significant effect of stress was found by the subject analysis ($F_1\,(1, 5) = 1.009, p = 0.361$) with only a marginally significant effect by the item analysis ($F_2\,(1, 5) = 5.288, p = 0.07$). The significant main effect of sentence type in both the subject analysis ($F_1\,(1, 5) = 4.204, p = 0.096$), and the item analysis ($F_2\,(1, 5) = 47.545, p = 0.001$), indicated that declarative question sentences (DQ) (*125 Hz*) had a higher F0 than the declarative assertion sentences (DA) (*118 Hz*) without considering stress location. The interaction between sentence type and stress type was not significant in both the subject analysis ($F_1\,(1, 5) = 4.121, p = 0.098$) and the item analysis ($F_2\,(1, 5) = 1.495, p = 0.276$). This non-significant interaction indicated that the obtained F0 effects on stress were due to sentential intonation rather than lexical differences. Therefore, F0 cannot distinguish the stressed syllables from the unstressed syllables regardless of sentence type. Overall these results indicate that speakers raised the F0 sentence finally regardless of stress location due to sentence type, specifically with the DQ intonation. The obtained F0 increase is due to sentence intonation. Therefore, the F0 is not a cue to stress location even in a highly marked context such as a question intonation.

4.3.1 Duration

Duration was measured on the first syllables in disyllabic nouns that contrasted on the first syllables in terms of stress. We treated the sentence type and stress type as within-subject variables as shown in Figure 11.

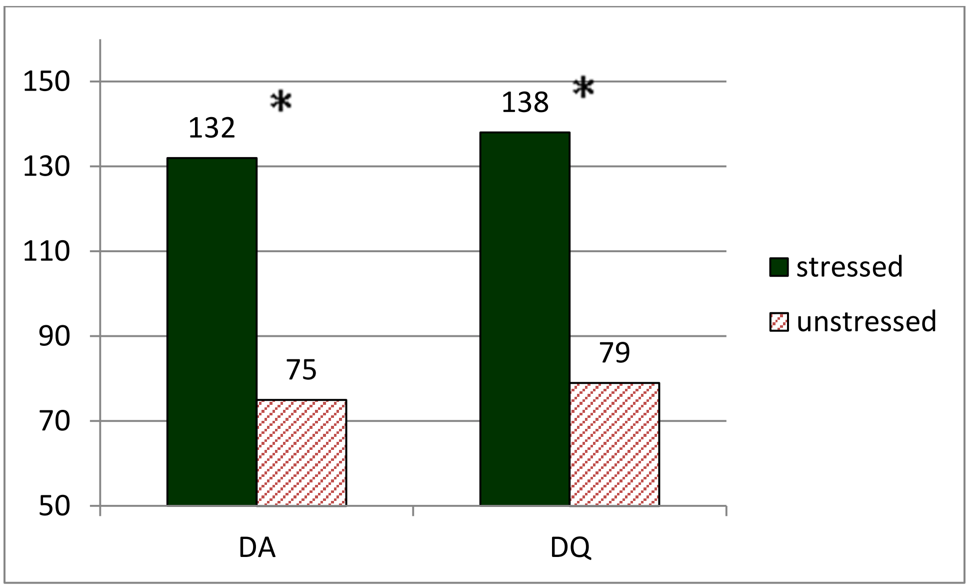

Figure 11 Average duration (ms): Dark bar shows data for the stimuli with first syllable stressed and the crossed bar shows data for the stimuli where the first syllable of the counterpart is not stressed. X axis shows declarative assertion (DA) versus declarative question (DQ) condition and Y axis is duration values in ms.

A strong significant main effect of stress type in both the subject analysis ($F_1\ (1, 5) = 101.737, p = 0.0001$) and the item analysis ($F_2\ (1, 5) = 35.813, p = 0.002$) indicated the stressed syllables *(135 ms)* had a significantly longer duration than the unstressed syllables *(77 ms)*. A significant effect of duration on stress type revealed that duration is a cue to stress location. A non-significant effect of sentence type for duration by the subject analysis ($F_1\ (1, 5) = 1.385, p =$

0.292) with a significant main effect of sentence type in the item analysis *(F_2 (1, 5) = 37.692, p = 0.002)* suggested that overall duration in the declarative question sentences (*109 ms*) was not significantly different from overall duration in the declarative assertion sentence (*104 ms*), without considering stress location.

In addition, the interaction between sentence type and stress location was not significant in both the subject analysis *(F_1 (1, 5) = 0.703, p = 0.44)* and the item analysis *(F_2 (1, 5) = 0.872, p = 0.393)*. No significant interaction indicated that the obtained duration effect was not due to sentential intonation, but due to lexical stress. In other words, even though the overall duration was different in DQ and DA sentences (only in the by item analysis), the non-significant interaction suggested that the changes in duration from the stressed to the unstressed syllables in DA sentences were the same as in DQ sentences. Therefore, the obtained duration effect is due to lexical stress.

4.3.2 Intensity

For the intensity analysis, we treated the sentence type and stress type as within subject variables as shown in Figure 12.

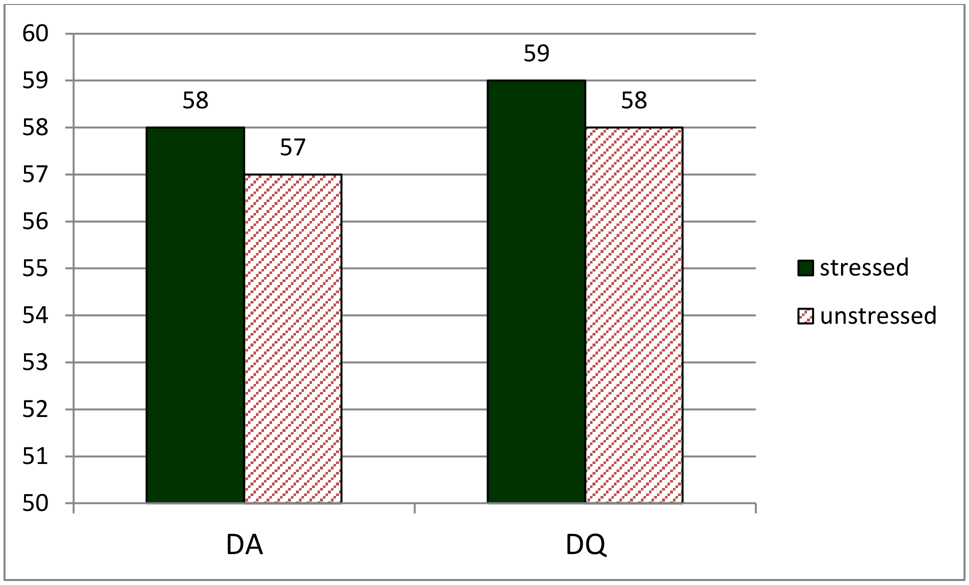

Figure 12 Average intensity (dB): Dark bar shows data for the stimuli with first syllable stressed and the crossed bar shows data for the stimuli where the first syllable of the counterpart is not stressed. X axis shows declarative assertion (DA) versus declarative question (DQ) condition and Y axis is intensity values in dB.

No significant main effect of stress type in either the subject analysis *(F_1 (1, 5) = 3.111, p = 0.138)* or the item analysis *(F_2 (1, 5) = 2.469, p = 0.177)* indicated the intensity in the stressed syllables *(59 dB)* was not significantly greater than in the unstressed syllables *(58 dB)* across sentence types. A non- significant effect of sentence type in both the subject analysis (F_1 *(1, 5) = 0.57, p = 0.484*) and the item analysis *(F_2 (1, 5) = 4.84, p = 0.079)* signified that overall intensity in DQ sentences (*59 dB*) was not stronger than in DA sentences (*58 dB)* across stress locations.

The interaction between the sentence type and stress location was also not significant in either the subject analysis *(F_1 (1, 5) = 0.002, p = 0.962)* or the item analysis *(F_2 (1, 5) = 0.034, p = 0.862).* The results indicated that the participants' behaviors in both the stressed syllables and the unstressed syllables were the same on DA and DQ sentential conditions.

4.4 Discussion

Experiment 3 examined the role of sentential intonation on lexical stress. In Experiment 1 and 2, there was no F0 role in distinguishing the stressed syllables from the unstressed syllables. In previous experiments (Experiment 1 and 2), the target words were arranged into a fixed sentence such as 'I will say ___ now.' In that case, all target words had the same sentential intonation. In Experiment 3, in order to investigate the role of F0, we chose declarative assertion (DA) and declarative question (DQ) sentence types. In DA and DQ sentential conditions, each target word occurred in the final position of the sentence, and the only difference was that F0 was rising for DQ sentences and F0 was level or falling for DA sentences. In Experiment 3, we only examined the CV-consistent stimuli group. From Experiment 2, we showed that the CV-consistent group produced systematic results in which duration was a strong cue, while intensity and F0 were not cues for stress location. Therefore, we used the CV-consistent group stimuli to examine the role of F0, duration and intensity in Experiment 3. The results showed the same pattern as in Experiment 2. Duration is a strong cue for differentiating the stress location, and intensity and F0 have no role. Even in a strong position of F0 (final sentential position), F0 still did not distinguish the stressed syllables from the unstressed syllables. We did find that our sentential manipulation was effective. We found that stimuli in DQ sentence contexts had a higher F0 than the stimuli in DA sentence contexts due to the fact that the declarative question sentences had a rising F0. Nevertheless, we found that even in DQ contexts, F0 does not provide a cue to lexical stress. Therefore, Uyghur patterns as a stress-accent language. We provide a summary for each parameter below.

In Experiment 3, F0 was measured on the first syllables of target words that were in DA (declarative assertion) and DQ (declarative question) sentences. We conclude that F0 does not distinguish the stressed syllables from the unstressed syllables in disyllabic nouns that had first-syllable stress contrasted. F0 in DQ was higher than DA conditions, suggesting that participants used rising intonation for declarative question sentences. However, no main effect of stress type and no interaction between stress type and sentence type suggested that F0 does not provide a cue for stress location. The obtained effect of F0 was due to the sentence intonation, but not due to lexical stress, because there was no significant interaction between sentence type and stress type overall.

We measured duration in the stressed syllables and the unstressed syllables of target words in different sentence types (DA versus DQ). From the duration analysis, the stressed syllables had a longer duration than the unstressed syllables across sentence types. There was no significant interaction with sentence type and stress type; therefore, the obtained duration effect in the stressed syllables versus the unstressed syllables was due to the lexical level rather than to sentential intonation.

We measured average intensity in the stressed syllables and the unstressed syllables in different sentence types (DA versus DQ). In terms of stress location, intensity of the stressed syllables was not greater than the unstressed syllables. There was no significant interaction with sentence type and stress. Intensity effects did not play a role either at the lexical level or for sentential intonation.

In conclusion, even in a strong position of F0 in declarative question sentences, neither intensity nor F0 was a stress cue; duration was a strong cue. The results confirmed the results from Experiment 2 in which disyllabic nouns were tested in a fixed condition. The results further suggest that F0 and intensity are not involved in lexical stress, while duration consistently marked the stressed syllable in Uyghur.

Chapter 5: Experiment 4 Acoustic Studies of Stress Patterns in Uyghur Learners

5.1 Experiment 4a an acoustic study of minimal pairs by Uyghur learners

5.1 Motivation and research questions

5.1.1 Motivation

Recently, lexical stress has been investigated not only among native speakers, but also among the learners (Zuraiq and Sereno, 2007; Lai, 2008). Zuraiq and Sereno (2007) examined not only English and Arabic stress pattern respectively, but also investigated the acquisition of English stress by Arabic native speakers whose second language is English. As a first step, Zuraiq and Sereno (2007) examined how English stress was produced by native English and non-native Arabic speakers. They used eight English minimal pairs that contrasted in terms of stress location and measured fundamental frequency, duration, intensity and vowel reduction. Zuraiq and Sereno (2007) found that English native speakers consistently used four parameters to distinguish the stressed syllable from the unstressed syllable. In other words, English native speakers used higher F0, longer duration, greater intensity, and reduced vowels in the stressed syllables than the unstressed syllables. However, Arabic speakers only used duration and intensity in a similar way to English native speakers, but they did not make similar use of vowel reduction. In terms of fundamental frequency, Arabic speakers used F0 to a larger extent compared to native English speakers. In order to know how their native language influences stress assignment in their second language, Zuraiq and Sereno (2007) also used six minimal pairs in Arabic. They found that Arabic speakers consistently used fundamental frequency, duration and intensity, but not vowel reduction in assigning stress in Arabic. They concluded that Arabic speakers had many similarities to English native speakers in using duration and intensity as cues to lexical stress, but that the Advanced Arabic speakers overused fundamental frequency in

English, in which the degree of difference in use of F0 was bigger than the native English speakers, and all Arabic speakers, regardless of proficiency levels, did not use vowel reduction.

In Uyghur, data from Experiment1 showed that duration and intensity provide cues for stress assignment. In English, fundamental frequency, duration and intensity are cues for assigning stress (Beckman, 1986; Fry, 1955). We predict that English speakers of Uyghur will use all three parameters if their native English language stress pattern influences their second language. We could also predict that if they have acquired Uyghur stress in their production, they might have a similar pattern to Uyghur native speakers. In other words, we want to examine how non-native Uyghur learners (who are native speakers of English) produce Uyghur stress, and how similar or different their pattern is from native Uyghur speakers.

5.1.2 Research questions and hypotheses

Our research question investigates the acquisition of the Uyghur lexical stress by native English speakers (learners of Uyghur):

Research question: Are the parameters that provide cues for accent location for native speakers of Uyghur used effectively by Uyghur learners when they produce Uyghur minimal pairs that contrast in stress?

We can divide the research question into the following:

(a) Do Uyghur learners use fundamental frequency as an acoustic cue for assigning stress in Uyghur (unlike Uyghur native speakers)?

(b) Do Uyghur learners use duration as an acoustic cue for assigning stress in Uyghur?

(c) Do Uyghur learners use intensity as an acoustic cue for assigning stress in Uyghur?

5.2 Methods

5.2.1 Stimuli

We used the minimal pair stimuli of Experiment 1.

5.2.2 Participants

Five (1F, 4M) native English speakers from the USA were recorded individually as shown in Table 15 below. Speakers are graduate students or self-employed and they learned Uyghur as a foreign language. The age range was from 24- 48 years old. All started studying Uyghur at a late age (AOA: 22-25 years)

Table 15 The Background information of Uyghur learners

participants	study Uyghur	stay in Xinjiang	AOA	other language	Chinese
1	5 years-grad	short visits	25	n/a	near-native
2	3 years-grad	7.5 months	22	Tajik	fair
3	informal 5 years	5 years	25	Kazakh, Uzbek	fair
4	2 years-grad	none	24	Spanish	none
5	2 years-grad	1 year (intensive)	22	Swedish, Japanese	good

AOA: Age of acquisition; none of them studied Uyghur as an undergraduate school

They were exposed to Uyghur informally via music, magazines, TV, traveling and Uyghur friends. All confirmed that they have slight accent in Uyghur. None of them acquired Uyghur before college or at the undergraduate level. Four of the five knew Chinese.

5.2.3 Procedure

Two participants were recorded in anechoic chamber with a Marantz PMD 671 solid-state recorder via an Electro-Voice ND767a microphone at the University of Kansas. Two other participants were recorded at the recording room of the Center for Languages of Central Asian

Regions (CELCAR) at Indiana University. For these two participants, we also used a Marantz PMD 671 solid-state recorder via an Electro-Voice ND767a microphone. The last participant was recorded in a quiet room, using a ZOOMH4ext handy recorder and a SHURE-SM10A microphone. All recordings were digitized at a sampling rate of 22.05 kHz except the last one, which was at 44 kHz. A total of 36 tokens were analyzed for each speaker. The total tokens are 180 (5 learners × 6 pairs × 2 words × 3 repetitions).

5.2.4 Measurements

Measurements were the same as in Experiment 1.

5.2.5 Statistics

We pre-tested the vocabulary, in order to determine whether the non-native speakers knew the words they produced. Participants read the vocabularies that were to be recorded and marked if they knew the words. After we recorded all stimuli, we did not analyze the recordings for the words that the participants did not know. When we removed one of the members in the paired words, we also removed the counterpart and all three repetitions as well. We analyzed only the stimuli that each participant knew. At the result, we threw out 90 tokens among 180 possible tokens from the five participants. Therefore, for these data, we only provide descriptive analyses. Due to great amount of missing data, none of the statistics reached significance. We believed the statistics did not have enough power because of missing data. For the data, we report values for each measure.

5.3 Results

5.3.1 Average F0

In Figure 13, the average F0 (absolute values across subjects) is provided for the two locations of lexical accent. English learners of Uyghur tended not to vary F0 due to syllable position in stress location in Uyghur. For example, in the initial stressed words, participants have an average F0 of *148Hz* on the first syllable and *147Hz* on the second syllable. In the final stressed words, participants showed no differences in F0 for first syllables *(150Hz)* as compared to the second syllables *(149Hz)*. Even though stress location was contrasted in the minimal pairs, the learners did not use F0 to indicate stress. Therefore, in terms of F0, the learners behaved very similarly to the native speakers of Uyghur who do not use F0 to indicate stress.

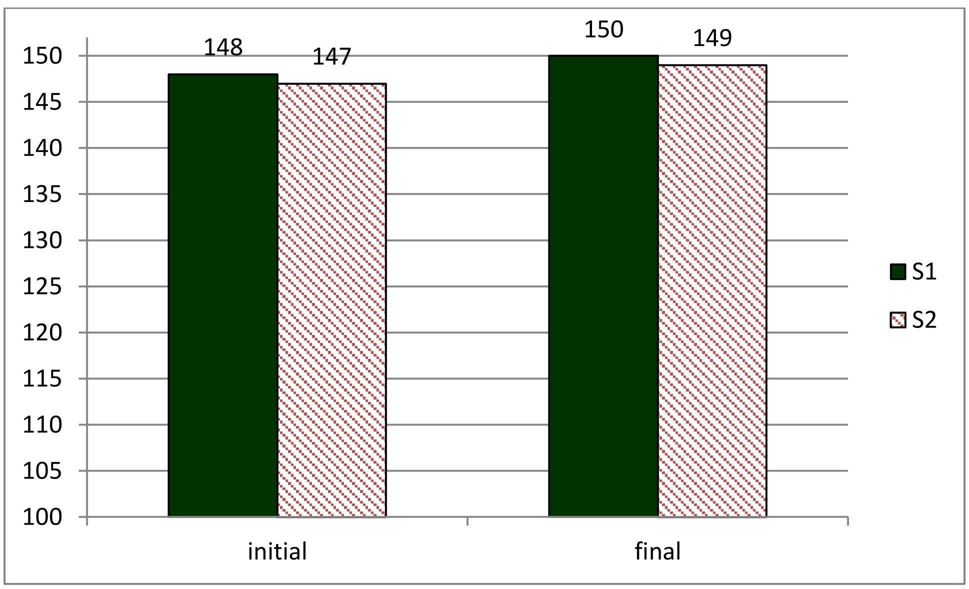

Figure 13 Average F0 (Hz): Dark bar shows data for the first syllable stressed stimuli and the crossed bar shows data for the second syllable stressed stimuli. X axis shows initial stress words versus final stress words and Y axis is F0 values in Hz.

5.3.2 Duration

In Figure 14, the durations (absolute values across subjects) are provided for each of the two locations of lexical accent. Both the initial and final stress words had a longer duration in the second syllables. When teaching Uyghur, stress is usually persistently taught as having final stress. If the learners were taught that the stress pattern in Uyghur is on the final position of the words, then they used the rule of stress for the duration. Non-native speakers did not use duration as a cue for producing stressed and unstressed syllables. Rather they assumed that stress was on the final position and increased the duration of that syllable for all stimuli.

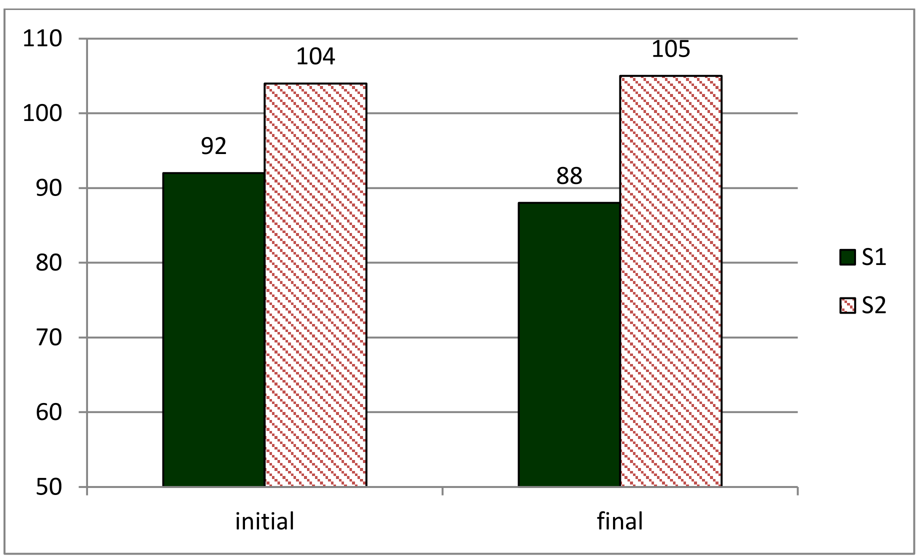

Figure 14 Average duration (ms): Dark bar shows data for the initial stress stimuli and the crossed bar shows data for the final stress stimuli. X axis shows initial stress words versus final stress words and Y axis is duration values in ms.

5.3.3 Intensity

In Figure 15, the average intensities (absolute values across subjects) are provided for the two locations of lexical accent. This showed the intensity was not a cue. English learners of Uyghur show few differences in intensity across syllable position or stress position.

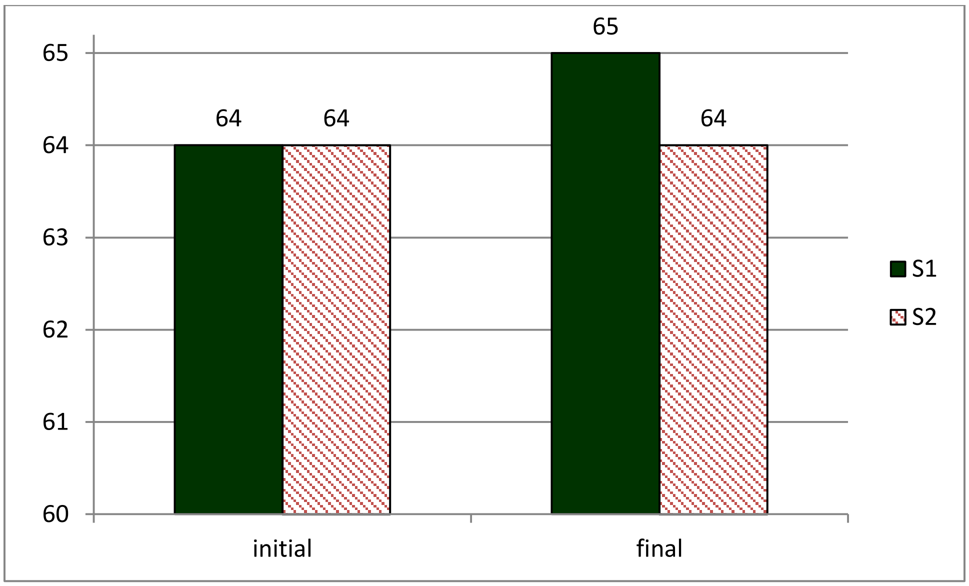

Figure 15 Average intensity (dB): Dark bar shows data for the initial stress stimuli and the crossed bar shows data for the final stress stimuli. X axis shows initial stress words versus final stress words and Y axis is intensity values in dB.

5.4 Discussion

In English learners of Uyghur producing minimal pairs in Uyghur, the pattern of leaners was neither like English nor like Uyghur. Their stress pattern on these minimal pairs was in between English and Uyghur. They did not use F0 and intensity cues at all. In terms of duration, they seemed to use the rule-based stress pattern, that is typically taught in the Uyghur classroom, for all words. Another possibility is that English learners of Uyghur used final lengthening instead of learning the rule-based stress pattern in Uyghur. Because both of these explanations would result in the same pattern of data (final lengthening), it is hard to make a solid conclusion.

In addition, the results cannot be directly compared to Experiment 1, which was about minimal pairs of Uyghur produced by native Uyghur speakers. In Experiment 4a, we only provided descriptive analyses due to missing data (i.e., the low frequency of some members of the minimal pairs). We concluded that non-native speakers did not systematically use any of the parameters in producing the stressed syllables versus unstressed syllables.

5.2 Experiment 4b: Disyllabic Noun Group Contrasting on Stress

5.2.1 Motivation and research question

5.2.1.1 Motivation

In Experiment 4a, we had 50% missing data because learners did not know the meaning of the words in one or both members in the Uyghur minimal pairs. In Experiment 4b, we used more representative and higher frequency stimuli to examine how learners acquire the stress pattern in Uyghur. The stimuli in Experiment 4b are relatively higher frequency compared to those of Experiment 4a; the majority of the participants knew their meanings. Finally, in Experiment 4a, the results were unclear since we could not answer the questions whether learners really did not use the all parameters or did the large amount of missing data cause the non-significant results? When we looked at Experiment 4 a, which was minimal pairs produced by English learners of Uyghur, the learners systematically used no acoustic parameters, and the patterns was neither like Uyghur nor like English. In order to know the acquisition pattern of stress by non-native speakers, we used the same stimuli from Experiment 2, which were disyllabic nouns contrasting on stress, to examine the English learners of Uyghur's production of more common Uyghur words.

5.2.1.2 Research Questions and Hypotheses

We used stimuli in Experiment 2 to examine learners stress patterns in Uyghur. Our research question investigates the acquisition of the Uyghur lexical stress by native English learners of Uyghur.

Research question 1: Are the parameters that provide cues for stress location for native speakers of Uyghur used effectively by Uyghur learners when they produce disyllabic words that contrast in stress?

We can divide the research question into the following:

(a) Do Uyghur learners use fundamental frequency as an acoustic cue for assigning stress in Uyghur (unlike Uyghur native speakers)?

(b) Do Uyghur learners use duration as an acoustic cue for assigning stress for Uyghur?

(c) Do Uyghur learners use intensity as an acoustic cue for assigning stress in Uyghur?

Hypothesis1: If Uyghur learners are influenced by English, they will use F0, duration, and intensity for differentiating the stressed syllables from the unstressed syllables in Uyghur.

Hypothesis 2: If they can separate Uyghur stress and English stress, they might exhibit the same patterns as Uyghur native speakers who use duration cues, but not F0 and intensity.

Research question 2: Are Uyghur learners sensitive to syllabic structure and consistency?

5.2.2 Methods

5.2.2.1 Stimuli

The stimuli were the same as in Experiment 2, which focused on disyllabic nouns that contrasted on the first syllables in terms of stress (see Table 8-10 in Experiment 2). The three groups included the CV-consistent, the CVC-consistent and CV-inconsistent stimuli groups.

5.2.2.2 Procedure

The procedure was the same in Experiment 2 using disyllabic noun contrasting in stress location.

5.2.2.3 Measurements

In this data set, we measure mean fundamental frequency, duration and mean intensity, which were identical to Experiment 2.

5.2.2.4 Participants

The participants were the same as Experiment 4a. Five native English speakers from the USA, who learned Uyghur as a foreign language, were recorded individually. A total of 108 tokens were analyzed for each speaker. The total tokens are 540 (5 learners × 18 pairs × 2 words × 3 repetitions).

5.2.2.5 Statistics

We used Repeated Measure ANOVA by subjects and by items. In the subject analysis, we averaged across all items' values in the CV-consistent, the CVC-consistent and the CV-inconsistent groups. Subject analyses showed the subjects' behaviors across items in the stressed and unstressed syllables. In the item analysis, we averaged across all subjects on each item. Item analyses indicated how each item's behavior changes across subjects in the different conditions.

First, we use a Repeated Measure (3x2) ANOVA, in which group has three levels (CV-consistent, CVC-consistent, and CV-inconsistent) and stress condition has two levels (first syllable-stressed and first syllable-unstressed), for all subjects and items separately. For these analyses, we treated groups (syllable structure or consistency) and stress condition as within subject variables for subject analysis; however for item analysis, we treated stress condition as a within subject variable, and groups (syllable structure or consistency) as between subject variables, respectively.

Secondly, we conducted the CV-consistent versus the CVC-consistent structure repeated Measure (2x2) ANOVAs in which syllable structure has two levels (the CV-consistent and the CVC-consistent), and stress condition has two levels (the stressed and unstressed syllables). The purpose of doing this analysis is to examine whether the syllable structure affects the stress patterns in the CV and the CVC stimuli.

In order to investigate consistency effects of perceiving stress location, we used Repeated Measure (2x2) ANOVAs in which consistency has two levels (CV- consistent versus CV- inconsistent) and stress has two levels (stressed syllables versus unstressed syllables). The purpose of doing this analysis is to examine whether consistency affects the stress pattern. In other words, we controlled the syllable structure and manipulated the perceived consistency.

5.2.3 Results

We measured the mean F0, duration, and the mean intensity on the vowels in the first syllable of the disyllabic words in the stressed and the unstressed conditions. We did not analyze two pairs of stimuli in the CV inconsistent group and one pair in the CVC group. The reason was that in Uyghur, high vowels occurring between voiceless consonants or preceding voiceless consonant are often devoiced. This devoicing also occurred in English learners of Uyghur. For example, the word *kiSHI* [kiʃi], 'person,' has stress on the second syllable and the first syllable was unstressed which was *ki*. The /i/ after the voiceless consonant [k] was devoiced. If one member of pairs was eliminated from the data, we also removed the counterpart (*kino*, 'movie') at the same time. The other example in CV inconsistent group was *SHIre* [ʃiɹɛ], 'desk,' and the counterpart was *shiPA* [ʃipa], 'cure.' For both of them, the first vowel /i/ was devoiced. The last pair was from the CVC group in which *PUTbol* [pʰutbol], 'football,' and its counterpart *putLASH* [pʰutlaʃ] 'trapping' were eliminated from the data analysis. The devoicing occurred across all Uyghur learners.

In addition, we pre-tested the participants to determine whether they knew the stimuli words or not. Based on the results, we removed the one item from the CVC-consistent group because none of the participants knew the word *SEPkün* [sɛpkyn], 'mole,' in Uyghur, and we also removed the counterpart *sepDASH* [sɛpdaʃ], 'fellow,' in the list even though they knew this word. Therefore, we had six pairs in the CV-consistent group, four pairs in the CVC-consistent group and four pairs in the CV-inconsistent group. At the result of pre-testing, 93 tokens were removed from a total of 540 tokens. Therefore, the missing data was 20.7%, much better than the data loss in Experiment 4a.

We had three groups: the CV-consistent, the CVC-consistent, and the CV-inconsistent stimuli. We show the results based on the F0, duration and intensity. We present the data in three sections as we mentioned previously. We first present overall results, then the syllable structure analysis, and finally the consistency results.

5.2.3.1 Average F0

As shown in Table 16, the average F0 (absolute value) was provided for the first syllables of the word pairs in three groups. The main effect of stress was not significant in both the subject analysis *(F_1 (1, 4) = 0.805, p = 0.42)* and the item analysis *(F_2 (1, 11) = 2.125, p = 0.173)*. There was no significant difference between the stressed syllables *(148 Hz)* and the unstressed syllables *(150 Hz)* regardless of stimuli groups. The results showed F0 cannot differentiate the stressed syllables from unstressed syllables. The main effect of group was not significant in either the subject analysis *(F_1 (2, 8) = 1.23, p = 0.342)* or the item analysis *(F_2 (2, 11) = 2.982, p = 0.217)*. This indicated that the patterns of three groups of stimuli were not different, regardless of stress location. The interaction between the group and stress type was not significant in either the subject analysis *(F_1 (2, 8) = 0.314, p = 0.739)* or the item analysis *(F_2 (2, 11) = 2.982, p = 0.092)*.

Table 16 F0 (Hz)Values and Standard Deviation among the Groups

Stress type	CV_con	CVC_con	CV_incon	overall AVG
stressed	144 (42)	151 (37)	148 (46)	148
unstressed	144 (41)	152 (49)	153 (43)	150
overall AVG	144	152	151	

Numbers in parentheses is standard deviation

For the F0 analysis, we concluded that F0 cannot distinguish the stressed syllables from the unstressed syllables across three groups for the non-native learners of Uyghur as shown in Figure 16. There were no consistency and syllable structure effects among the non-native speakers, because the main effect of group was not significant. Therefore, we can claim that F0 is not a cue to indicate stress location in non-native speakers who have relative high advanced level, a pattern very similar to Uyghur native speakers. Uyghur learners also were not sensitive to either syllable structure or consistency in terms of F0.

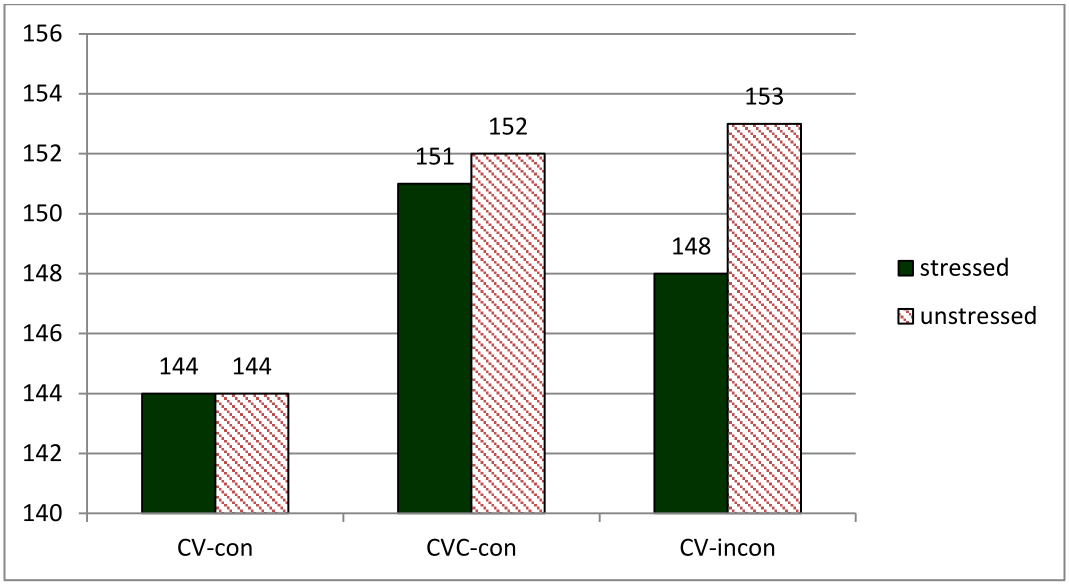

Figure 16 Average F0 (Hz): Dark bar shows data for the stimuli with first syllable stressed and the crossed bar shows data for the stimuli where the first syllable of the counterpart is not stressed. X axis shows three groups and Y axis isF0 values in Hz.

5.2.3.2 Duration

(1) Overall results

As shown in Table 17 below, the average duration (absolute value) was provided for the first syllables of the word pairs in the three stimulus groups. The main effect of stress was significant in both the subject analysis (F_1 $(1, 4) = 24.919, p = 0.008$) and the item analysis (F_2 $(1, 11) = 5.177, p = 0.044$). The stressed syllables (*92ms*) had a longer duration than the unstressed syllables (*76ms)* regardless of group. The results showed that duration can differentiate the stressed syllables from the unstressed syllables. The main effect of group was significant by the subject analysis (F_1 $(2, 8) = 10.131, p = 0.006$), but not by the item analysis (F_2 $(2, 11) = 1.413$,

$p = 0.257$). This indicated that the patterns of three groups were different in the subject analysis. In other words, groups (*101ms* for CV-consistent, *67ms* for CVC, and *85ms* for CV-inconsistent) were slightly different from each other. The interaction between the group and stress type was also significant in the subject analysis *(F_1 (2, 8) = 11.653, $p = 0.004$)*, but not by the item analysis *(F_2 (2, 11) = 1.541, $p = 0.257$)*.

Table 17 Duration (ms) Values and Standard Deviation among the Groups in NNS

Stress type	CV_con	CVC_con	CV_incon	overall AVG
stressed	105 (22)	69 (8)	102 (16)	92
unstressed	97 (11)	64 (9)	68 (4)	76
overall AVG	101	67	85	

Numbers in the parentheses are standard deviation; NNS is non-native speakers

We also ran a paired t-test for the stressed syllable versus the unstressed syllables for each group in terms of duration. In the CV-consistent group, no significant difference was found between the stressed syllables *(105ms)* versus the unstressed syllables *(97ms)* by the subject analysis (*t_1 (4) = 1.572, $p = 0.191$)* and by the item analysis (*t_2 (5) = 0.726, $p = 0.501$)*. In the CVC-consistent group, no significant difference between the stressed syllables (*69ms*) versus the unstressed syllables analysis (*67ms*) was found for the subject analysis *(t_1 (4) = 1.467, $p = 0.216$)* or the item analysis *(t_2 (3) = 1.279, $p = 0.291$)*. In the CV-inconsistent group, there were significant differences between the stressed syllables (*102ms*) versus the unstressed syllables *(68ms)* by the subject analysis *(t_1 (4) = 5.783, $p = 0.004$)* and significant by the item analysis *(t_2 (3) = 3.12, $p = 0.052$)* as shown in Figure 17.

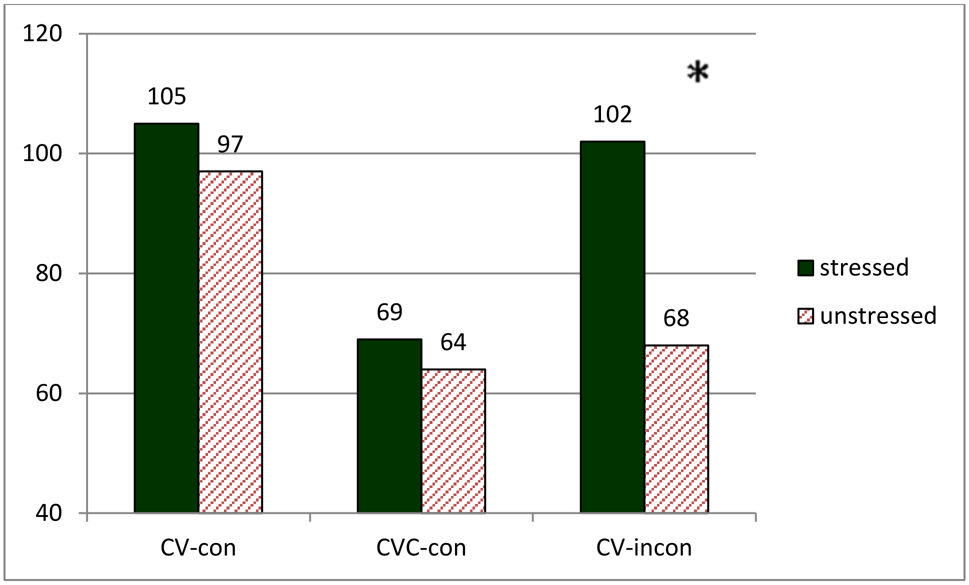

Figure 17Average duration (ms): Dark bar shows data for the stimuli with first syllable stressed and the crossed bar shows data for the stimuli where the first syllable of the counterpart is not stressed. X axis shows three groups and Y axis is duration in milliseconds.

From the overall analysis of duration, we obtained a significant stress effect both in subject and item analyses. Moreover, a paired t-test showed that duration does matter most for the CV-inconsistent group. We also conducted syllable structure and consistency analyses.

(2) Syllable structure

In order to examine the syllable structure effects, we contrasted the CV-consistent and CVC-consistent groups and used a two-way Repeated Measure ANOVA. The main effect of stress-type was marginally significant by the subject analysis (F_1 *(1, 4) = 0.6.44, p = 0.064*), but not the item analysis (F_2 *(1, 8) = 0.786, p = 0.406*). This showed that *87ms* for stressed syllables were not significantly longer than the *81ms* for unstressed syllables regardless of syllable type. This seems like duration does not provide information about stress location in the consistently perceived syllables with different syllable structure types in disyllabic nouns in non-native speakers. However, when we look at the trend, we find that the stressed syllables are slightly longer than the unstressed syllables in CV-consistent stimuli, but less different in the CVC-stimuli. It is possible that the number of speakers and number of stimuli cause more variability. Moreover, the main effect of syllable type was significant by the subject analysis (F_1 *(1, 4) = 10.646, p = 0.031*) and marginally significant by the item analysis (F_2 *(1, 8) = 3.54, p = 0.097)*. The CVC syllables (*67ms*) were shorter than the CV syllable structure (*101ms*) regardless of stress type. The interaction between stress type and syllable type was not significant in either the subject analysis *(F_1 (1, 4) = 0.2, p = 0.678)* or the item analysis *(F_2 (1, 8) = 0.078, p = 0.787).* This indicated that non-native learners of Uyghur were not sensitive to syllable structure or stress types.

From the syllable structure analysis, we found that different syllable structure (CV-consistent versus CVC-consistent) have different patterns in duration. Duration changes according to syllable type.

(3) Consistency effects

In order to investigate the consistency, we used CV-consistent and CV-inconsistent stimuli and used a two-way Repeated Measure ANOVA. The main effect of stress-type was significant in both the subject analysis *(F_1 (1, 4) = 19.657, p = 0.011)* and the item analysis *(F_2 (1, 8) = 5.145, p = 0.053)*. The stressed syllables *(104 ms)* were significantly longer than the unstressed syllables *(83ms)* regardless of consistency. This further supported that claim that duration does provide information about stress location in the same syllable structure with different levels of consistency for disyllabic nouns. Moreover, the main effect of consistency was significant by the subject analysis *(F_1 (1, 4) = 11.666, p = 0.027)*, but not by the item analysis *(F_2 (1, 8) =0. 633, p = 0.449)*. Duration of CV-consistent syllables (*101ms*) was not significantly different from the

CV-inconsistent syllables (*85ms*) regardless of stress type. The interaction between stress type and consistency was significant in the subject analysis *(F_1 (1, 4) = 20.377, p = 0.011)*, but not in the item analysis *(F_2 (1, 8) = 1.506, p = 0.255)*. There were few effects of consistency.

From the results of the overall analysis, syllable structure analysis and consistency analysis, we concluded that duration can differentiate stressed syllables from unstressed syllables in the Uyghur learners' productions. There were neither syllable structure effects nor consistency effects in terms of duration in non-native speakers. The strongest duration differences came from the CV-inconsistent stimuli.

5.2.3.3 Intensity

As shown in Table 18, the average intensity (absolute value) was provided for the first syllables of the word pairs in three groups. The main effect of stress was not significant in both the subject analysis *(F_1 (1, 4) = 1.901, p = 0.24)* and the item analysis *(F_2 (1, 11) = 0.104, p = 0.753)*. No significant difference was found between the stressed syllable (*65dB)* and the unstressed syllables (*64dB)* regardless of stimuli groups. The results show that the role of intensity does not seem to differentiate the stressed syllables from the unstressed syllables. The main effects of groups were not significant in both the subject analysis *(F_1 (2, 8) = 2.064, p = 0.189)*, and the item analysis *(F_2 (2, 11) = 1.172, p = 0.346)*. This indicated that the patterns of three groups were not different regardless of stress location. The interaction between the group and stress type was not significant in either by the subject analysis *(F_1 (2, 8) = 0.4, p = 0.683)* or by the item analysis *(F_2 (2, 11) = 1.172, p = 0.346)*. From the overall analysis of intensity, we obtained that the group effects were not significant, and the stress type effect was not significant.

Table 18 Intensity (dB) Values and Standard Deviation among the Groups

Stress type	CV_con	CVC_con	CV_incon	overall AVG
stressed	65 (8)	65 (9)	63 (8)	65
unstressed	65 (8)	64 (8)	62 (9)	64
overall AVG	65	65	63	

Numbers in parentheses indicate standard deviations

The result showed that intensity cannot differentiate the stressed syllables from the unstressed syllables in both subject and item analyses. The overall main effects of stress were not significant across syllable structure and consistency analyses. Intensity was not a cue to stress location in non-native learners of Uyghur.

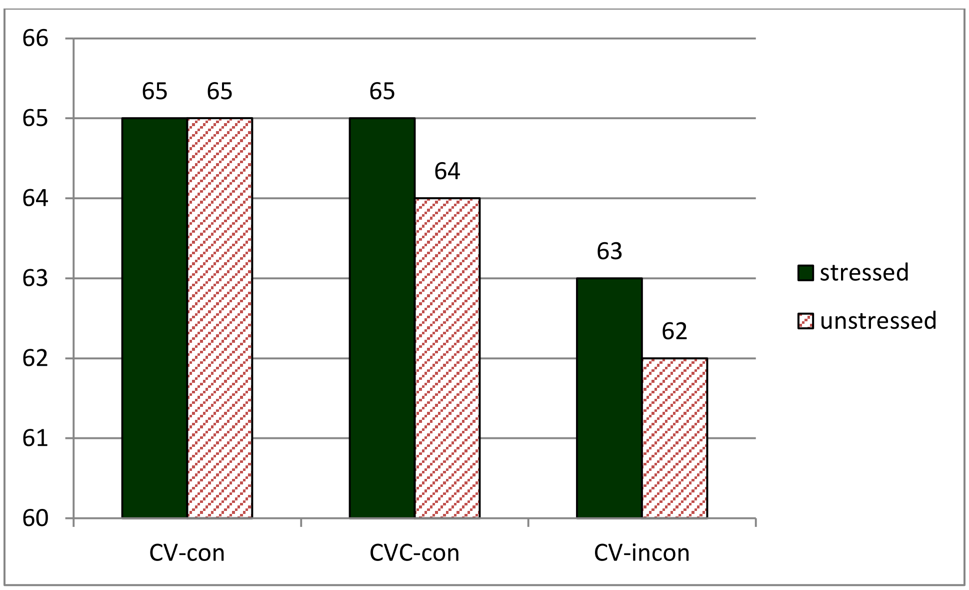

Figure 18 Average intensity (dB): Dark bar shows data for the stimuli with first syllable stressed and the crossed bar shows data for the stimuli where the first syllable of the counterpart is not stressed. X axis shows three groups and Y axis is intensity in dB.

5.2.4 Discussion

We used disyllabic nouns that contrasted on the first syllable to examine the acquisition of stress patterns by the non-native speakers who are learners of Uyghur. Five non-native speakers, who had more than 4 semesters learning Uyghur, and had stayed at least one year in the immersion of the Uyghur community (four out of learners five learners), produced the target words in a fixed sentence. Similar to the native speakers' data, we measured fundamental frequency, duration and intensity on the vowels that were in stressed and unstressed syllables.

Non-native speakers did not use F0 and intensity, which were similar to native speakers. But the result about the duration was somehow mixed. Native speakers used duration as a cue for stress location. A significant main effect of duration in the overall analysis indicated that duration was a cue for stress location also among the non-native speakers. However, in each group, when we compared the stressed syllables with the unstressed syllables, only CV-inconsistent showed significant results for the non-native speakers, while in native speakers all groups showed significant differences between the stressed syllables and the unstressed syllables. We believe that CV-inconsistent contributed the most to the duration significance in the overall analysis for non-native speakers. If we analyzed only CV-consistent versus CVC-consistent stimuli, duration did not show stress cues for the learners. It seems to be more lexically conditioned. Therefore, non-native speakers did not use F0 and intensity as cues for differentiating the stressed syllables from the unstressed syllables, while duration was a weaker cue for stress location among non-native speakers who are learners of Uyghur.

Chapter 6 General Discussion and Conclusion

6.1 Statement about the research

This research provides an acoustic analysis of lexical stress patterns in Uyghur. The basic research question was which acoustic parameters were strongly correlated to stress location in Uyghur. Usually, in a production study, participants produce minimal pairs in isolation or a fixed sentence. The fundamental frequency, duration and intensity are measured for vowels contrasting stressed and unstressed conditions. Based on a large number of studies on a number of languages, duration, intensity and fundamental frequency are often the critical acoustic cues for distinguishing the stressed syllables from the unstressed syllables. Based on the selected use of these acoustic cues, languages can be categorized as stress-accent, tone-accent or pitch-accent languages.

In the case of Uyghur, the debate has focused on whether Uyghur is a pitch-accent language similar to Turkish, or whether Uyghur is a stress-accent language similar to English. In order to determine the categorization of accent patterns of Uyghur, we conducted three experiments to examine the stress patterns in Uyghur by native speakers in a production study and then we conducted one experiment to examine the stress patterns in Uyghur learners.

6.1.1 Experiments with native speakers

We first conducted three experiments to examine acoustic correlates of stress in Uyghur. Experiments 1 and 2 focused on minimal pairs and disyllabic nouns in a fixed sentence condition. The two experiments showed which acoustic parameters provided cues for stress location. Experiment 3 focused on the interaction between sentential intonation and lexical stress.

Experiment 3 answered the question whether F0 plays a role in lexical stress when F0 is also used to indicate the intonation.

In Experiment 1, we used six minimal pairs in a fixed sentence condition and measured fundamental frequency, duration and intensity on the vowels that were in stressed and unstressed syllables. Using ratio analysis as Beckman (1986) used, we found that duration and intensity were cues for distinguishing the stressed syllables from unstressed syllables. F0 was not a stress cue.

In Experiment 2, we used disyllabic nouns that contrasted on the first syllables in terms of stress location. In addition, we manipulated the syllable structure and the consistency of stress assignment. Consequently, we had CV-consistent, CVC-consistent and CV-inconsistent stimuli groups. Similar to Experiment 1, we also measured fundamental frequency, duration and intensity on the vowels. We found that duration was a strong cue for stress location; F0 and intensity were not cues across the three groups of stimuli for distinguishing the stressed syllables from the unstressed syllables. When duration was a cue, participants were sensitive to syllable structure of the stimuli. In terms of F0 and intensity, participants were not sensitive to either syllable structure or consistency.

Concerning the research question about the role of the syllable structure, we obtained syllable structure effects only in duration. Overall, vowels in open syllables were longer than in closed syllables. The vowels in the CV-consistent group were also longer than the vowels in the CVC-consistent group in the stressed position. Interestingly, even CVC syllable structure is more complex than CV syllable structure, but CV syllable structure had stronger duration effects in terms of stress.

For the research question about whether the consistent stress assignment plays a role in stress location, there was no consistency effect observed. This indicated that even though we asked participants where the stress was located, and they had different consistency levels in assigning the stress in their productions, the speakers produced a similar pattern in consistent and inconsistent stimuli groups. While perceptually listeners are ambiguous about where stress is located, in production they consistently indicate stress using duration cues.

In Experiment 3, we used declarative assertion and declarative question type sentences for the stimuli, because we could control the position of the target words in the final position of the sentences where F0 differed in two different sentence types. In Experiment 2, we found that the CV-consistent stimuli produced more systematic results than the other groups, and therefore, we used the CV-consistent group stimuli to examine the interaction between sentential intonation and lexical stress. We measured the same parameters, such as fundamental frequency, duration and intensity, as in Experiment 1 and 2. We found that duration, but not intensity and F0, was a cue for stress location. No significant interaction between the stress type and sentential intonation indicated that the obtained duration effects belonged to lexical stress rather than sentential intonation.

From the overall results, we concluded that duration is a strong acoustic cue for stress location in Uyghur. F0 and intensity did not have any role in distinguishing the stressed syllables from unstressed syllables. Uyghur did not use pitch to distinguish the stressed syllables from the unstressed syllables, and therefore, Uyghur patterns very similarly to many stress-accent languages and patterns very differently from many pitch-accent languages.

6.1.1.1 Comparison with previous experiments on Turkic languages

Our results have some similarity with previous research (Liang and Zhang, 2008; Jiang et al, 2010; Dobrovolsky, 1999). First, Liang and Zhang (2008) examined the accent pattern in Uyghur using 16 disyllabic words and five newly created non-words. All stimuli were produced in isolation by 10 (5F, 5M) native Uyghur speakers. They measured F0 contour, duration and peak intensity on the syllables. In their research, accent pattern was not examined rather, the 16 disyllabic words included four syllable structures (e.g. CVCV, CVCVC, CVCCVC and CVCCV), and stress was always on the final syllable. They compared the first syllables to the second syllables with absolute values for each acoustic parameter. Across all four syllable groups, duration was longer in the second syllables than the first syllables. F0 contour was measured every 10ms and they found that in both first and second syllables F0 contour was dropping and F0 was higher at the initial position of the second syllables than the final position of the first syllables. Intensity had inconsistent results in which it was the same in the first and second syllables for some items while it was stronger in the second syllables compared to the first syllables for other times. Liang and Zhang (2008) concluded that duration could distinguish the stressed syllables from the unstressed syllables. The duration effects were the same across the four different syllable structures. They concluded that syllable structure did not influence stress pattern. However, the observed duration effects could be due to phrase final lengthening since they compared the second syllables to first syllables in terms of duration. Unlike their methods, we used first syllables that contrasted in terms of stress in disyllabic nouns in order to avoid the confounding effect of the final lengthening of the final syllables and the possible rising F0 in the final position.

Secondly, Jiang et al (2010) also investigated the stress pattern in Uyghur in disyllabic and trisyllabic words. In disyllabic words, they measured F0 at the vowel mid-point (the average of maximum and minimum F0s), duration and total intensity on the syllables for 20 disyllabic words in Uyghur. They also assumed that second syllables (final syllables) were stressed. The stimuli were produced in isolation by eight (4F, 4M) native Uyghur speakers. Unlike Liang and Zhang (2008), Jiang et al. provided only descriptive analyses with ratios (second syllable/first syllable) and did not provide any statistics. They found that in terms of the duration and intensity, the proportion of the second syllables' parameters was greater than the first syllables' parameters, but not for F0. Therefore, they concluded that only duration and total intensity could differentiate the stressed syllables from the unstressed syllables. They also did not control the syllable structure and it was problematic since they were measuring the duration and total intensity based on the syllable, and the stimuli included all combinations of V, CV, and CVC syllable structures.

Both studies focused on disyllabic words, and found that duration could distinguish the stressed syllables from the unstressed syllables. In terms of duration, their findings were similar to ours; our results also showed that duration was a cue for stress location. Unlike their results, the present data show that intensity and F0 did not play a role in stress locations. However, for intensity, Liang and Zhang (2008) used peak intensity, but Jiang et al (2010) used total intensity. It is possible that Jiang et al's use of total intensity covered more intensity information than Liang and Zhang's peak intensity. Unlike previous studies, the present data contrast stressed and unstressed syllables, using minimal pairs or minimal syllables. We also pre-tested the perception of stress location, and controlled the syllable structure and the consistency. We used average F0, duration, average intensity and we found only duration was a strong cue, while intensity depended on the lexical structure, and F0 was not a stress cue.

Unlike Levi's (2005) findings that Turkish used F0 as the only cue for assigning stress location, Uyghur used duration for assigning stress location, not F0 and intensity. Perhaps the reason for this difference stems from Levi's (2005) use of inflectional-level near-minimal pairs instead of word-level minimal pairs. These differences may be the result of the number of syllables used in the research. In the Turkish study, the words were at least trisyllabic words instead of disyllabic words; therefore, secondary stress might be a confounding factor in the results. In the absolute value comparison, Levi (2005) found that duration and intensity were able to distinguish the stressed syllable in Turkish from the unstressed syllables, but in Linear Discriminant Analysis (LDA), the effects of duration and intensity disappeared when these effects were compared to the effect of F0.

Like our stimuli, Dobrovolsky (1999) also examined disyllabic nouns in Chuvash, which is one branch of Turkic languages. Instead of using disyllabic nouns that contrast stress on the first syllable as we did, he controlled the disyllabic words in four word-stress groups: full-full (first syllable is full vowel and second syllable is also full vowel), full-reduced, reduced-full, and reduced-reduced disyllables. By measuring duration and intensity, Dobrovolsky found that the duration ratios and total amplitudes were significant cues within each word-stress class except for the reduced-reduced group, which was considered as a without stress group. Even though he used the different stress groups, the result confirmed that duration and intensity were the cues for Chuvash. Unlike our results, Chuvash also used intensity, but the role of intensity was not as strong as duration according to Dobrovolsky (1999).

In conclusion for the Turkic language studies, duration was a solid cue for Uyghur, Chuvash, and it could be a cue for Turkish too if we focused on the results with absolute values. However, the role of intensity was mixed across studies, and it depended on the language and measurements. F0 was not tested in Chuvash, but it was a cue for two studies (Uyghur by Liang and Zhang, 2008; Turkish by Levi, 2005). But our results and the results from another study (Jiang et al 2010) in Uyghur did not show the F0 effect. The inconsistent results may not be due to language, but more to methodology in this case.

6.1.1.2 The role of acoustic parameters

In our experiment, we obtained a duration effect that can be a stress cue. There are two possibilities about the duration effect. First, duration is a solid cue for stress-accent languages. Duration, being a stress cue in Uyghur, was confirmed by many studies (Fry, 1958 and Beckman, 1986, for English; Sluijter and Van Heuven, 1966a, 1996b, for English and Dutch; Sereno and Jongman, 1995, for English; Nash, 2005, for Sinhala verbs; Dobrovolsky, 1999, for Chuvash; Gordon, 2004, for Chickasaw, among others). Similar to the above studies, the stressed syllables had longer duration than the unstressed syllables.

Secondly, the present results clearly show that F0 was not a cue for Uyghur in terms of stress location. The finding that F0 effect is not being a stress cue differentiates Uyghur from both English and Japanese. The role of F0 has been debated for a while. Lindstrom and Remijsen (2005) proposed a no-pitch language by examining the Kuot language. Uyghur does not use the F0 to distinguish the stressed syllable from the unstressed syllables. Therefore, the results in Uyghur, which does not use F0, were similar to the results in Kuot. We propose that Uyghur does not pattern like a pitch-accent language, but rather like a stress-accent language that uses duration to cue stress.

6.1.1.3 Notes about the duration effects

In this section we examined whether the effects obtained in Uyghur may be confounded with other factors such as final lengthening, long vowels and the borrowed lexicon which have heavy syllables or stress. Firstly, we could safely eliminate final lengthening as an explanation of the duration effect in Experiment 2. We used the first syllable comparison instead of second or final syllable in order to avoid final lengthening. Therefore, the obtained duration effect in Experiment 2 could not be a final lengthening effect.

A second possible explanation is that the obtained duration effect could be due to the existence of long vowels or vowel lengthening. Here, we discuss the vowel lengthening or vowel length issues with words of Turkic, Arabic, Persian and Russian origin.

(1) *Turkic-originated words in our experiments*

Old Turkic and most Middle Turkic contrasted vowel length (Dwyer 2000, 2002) but Modern Standard Uyghur does distinguish vowel length. Even though these minimal pairs reflected vowel length distinctions in Old Turkic, Modern Standard Uyghur and many modern Turkic languages lost the vowel length distinction. Table 19 below shows how only one variety of modern Uyghur, in Kelpin, preserves Common Turkic vowel length distinctions (as aspiration, Dwyer, 2000), whereas modern Uyghur has lost these distinctions.

Table 19 Common Turkic and Modern Standard Uyghur (Dwyer 2000)

Common Turkic	Kälpin Uyghur	Standard Uyghur	gloss
o:t	ot^h	ot^h	‘fire’
ot	o^ht^h	ot^h	‘grass’
a:t	at^h	at^h	‘name’
at	a^ht	at^h	‘horse’
ɛ:t-(verb)	εt^h	εt^h	‘to do’
ɛt	$\varepsilon^h t^h$	εt^h	‘meat’
ø:t	$\o t^h$	$\o t^h$	‘gallbladder’
øt-	$\o^h t^h$	$\o t^h$	‘to pass’
a:q	aq^h	aq^h	‘white’
aq-	aq^h	aq^h	‘to flow’

Table 19 shows that the original long vowel in Common Turkic became a short vowel in Modern Standard Uyghur.

In terms of vowel length, there are two types of long vowels. One is primary long vowels, which exist in Old Turkic and most in Middle Turkic. The other one is secondary long vowels, which are caused by compensatory lengthening both diachronically or synchronically. In Modern Uyghur, /r/ was dropped in final coda position. For example, *bazaar* ‘market’ is usually pronounced as *baza:*. However, in writing the /r/ will show up. Some of them already existed diachronically. For example *ete* ‘tomorrow’ used to be *erte*. It is the same written and spoken as *ete*. The first ‘e’ is pronounced longer.

In our Experiment 1 and 2, the Old Turkic-originated words in our stimuli had long vowels in Old Turkic words, but now there are no long vowels in Modern Standard Uyghur as shown in Table 20 below (for full citation see appendix B).

Table 20 Turkic Stimuli Comparisons with Old Turkic Source (verbatim quotes from Clauson 1972)

Modern Uyghur	Gloss in Modern Uyghur	Early Turkic form	Gloss and early Turkic cognates
Acha	elder sister	eche:	'one's mother's younger sister; one's own elder sister'; Xak[4].: 'elder sister;'Chagh.: ece 'an elderly woman'
aRA	between	ara: (ʔa:ra)	'between' ; Xak.: *ara:* 'in the middle of things;' Chagh.: *ara*, 'middle, center'
BAla	child	bala:	'young bird, nestling'; Xak.: 'a nestling' metaphor for 'a young of any predatory animal;' *balu:* and *bala:* 'a helper for a man in his work' (Kash.); Chagh.: 'a young of animal'
töSHÜK	hole	teshük /teshik	'hole'; from *tesh*-; -lit. 'pierced' Xak.: 'raptured'; 'a glutton;' Chagh.: 'hole'
TÖshük	kitchen	unknown	Possibly related to *tütün* 'smoke'
daDA	father	dede:	Oghuz XI. 'father;'
putLASH	tripping	bu:t (bu:d) put 'foot'+verblizer -LA- + Nominilizer -(I)sh	'the thigh'; Xak.: *bu:t*, 'the thigh'; Chagh.: *but* 'the leg from thigh to the toes';
TOQmaq	stick	tokı:mak	from *tokı:*- 'hit, knock;' 'club, mallet;' Xak.: 'mallet;' Chagh.: *tokmak* 'well-known implement used to drive in tent pegs'
toqQUZ	nine	tokku:z	'nine;' Xak.: *toku:z*; Chagh.: *tokuz*;
soQA	plough	soku: (sokghu:)	from *sok*- 'beat, crush;' Xak.: soku: 'a mortar;' Chagh.: *sokku:* 'a large wooden mortar'
kiSHI	person	kishi:	'man, person, human being;' Xak.: 'a man, mankind'; Chagh.: 'man' in general, female and male; 'a man' in singular.
TOxu	hen	taki:ghu:	'a domestic fowl'; Xak.: *taka:ghu* 'cocks and domestic fowl'; Chagh.: *taghuk* or *taxuk* 'a bird'
toPA	dirt, soil	tüpi:	'high wind'; 'high wind carrying snow or dust'; Xak.: 'a high wind'

Only the words that had etymology source are listed

Table 20 shows that, the listed stimuli in Experiment 1 and 2 (except *sepkün*, and *töshük*, which did not have long vowels) had long vowels in Old Turkic words. In the recent usage in the Chaghatay language, the long vowels were shortened. For example, the long vowel in *ara:* in Old Turkic words became a short one in the Ghaghatay language and changed into *ara*. Especially in Experiment 2, we focused on the first syllables of the disyllabic words and the majority of long vowels occurred in the second syllables. Therefore, even though long vowels

[4] Xak. represents the Xakani language (middle Turkic from XI centuries); Chagh. represents the Chaghatay language. Kash. Kashgari, who wrote *the Diwani Lughati'l Turk*; Uyg VIII, a Uyghur language in VIII century;; Uzb represents Uzbek, SC language. K represents two /k/ and /q/ which depends on back vowel for /q/ and front vowel /k/

existed potentially, this could not influence the results. In addition, if long vowels occurred in the first syllables such as *bu:d* < *put*, the long vowel was shortened and changed into a short vowel in the Chaghatay language or in Modern Uyghur. The primary long vowels were shortened in Modern Uyghur. Moreover, Osmanov (2009) extremely claimed that even in Old Turkic words the vowel length was questionable. He investigated the Old Turkic texts (*Diwan Lugheti-i Turk*, The Compendium of Turkic dialects) in which the writing system differentiated the long vowels from short counterparts by using extra diacritics on the short version of the vowels. He found that there were not many existing long vowels (of 600 hundred stem-based words, only 78 items showed long vowels and in examining 17 words with their derivation and inflectional version, only the stem-based words showed long vowels but the derivational and inflectional version did not show long vowels) and even when there were some long vowels, they were not consistent throughout the Old Turkic Texts. From these studies we conclude that at least in Modern Standard Uyghur, there is no vowel length distinction. Therefore, the long vowels in Old Turkic words did not influence our analyses.

When we discussed the duration effects that correlated to stress location, we observed that the diachronic compensatory lengthening in Uyghur may be confounded with the duration effects. The compensatory lengthening indicated that some segments (and letters) in coda position were dropped historically. There were two stimuli that possibly had diachronic compensatory lengthening: (1) *TOQmaq* ‘stuck’ in CVC-consistent group was from Chagh. *tokmak* < Old Turkic *tokı:mak*. It is possible the /ɪ:/ vowel in *tokı:mak* in Old Turkic was dropped in the Chaghatay language and changed into *tokmak* with potential lengthening. We tested the disyllabic nouns that shared the same first syllable (*toqQUZ*, ‘nine’), and found there was no duration difference between the two syllables (t =1.56, p = 0.18). The diachronic compensatory lengthening did not exist in this pair. (2) *TOxu* 'hen' in the CV-inconsistent group was from Chagh. *taxuk* or *tavuk* < Old Turkic *taka:ghu*. *TOxu* in Modern Uygur could have potential lengthening if it was from *taka:ghu* or *tavuk*, which had some phonetic changes. The example *to* in *TOxu* was a long vowel after diachronic compensatory lengthening. When we chose the counterpart, which was *toPA*, and compared the durations of the first syllables of *TOxu* and *toPA*, we found there was a significant difference between them ($t = 3.09$, $p = 0.02$). However, the words (*TOxu* vs. *toPA*) were in the CV-inconsistent group which means participants' assignment of stress had variability.

(2) Russian borrowed words in our experiments

Turkic languages have exceptions in stress location in which some classes of words have initial stress rather than final, regular stress (Kodzasov, 2003). With the long term language contact, Uyghur borrowed many words from Arabic, Persian, and Russian. The borrowing could be one source of having exceptional stress in Turkic languages including Uyghur (Kodzasov, 2003). We listed the words in Experiment 2 (Experiment 1 did not include Russian loans), which were disyllabic nouns contrasted on the first syllables and these words were from Russian as shown in Table 21 below. In Russian, stress is 'free' in which it could be on the first syllable or medial or final position (Wade, 1992:14). Hamilton (1980) indicated Russian does not have vowel length distinctions; however the vowels are longer under the stress positions than the unstressed positions. In other words, vowel length is not phonemic in Russian, similar to English. Russian uses duration, intensity and vowel reduction as cues for the stress location (Hamilton, 1980; Kuznetsova, 2006).

Table 21 Russian Borrowings with Modern Uyghur Perceived Stress (Katzner 1994)

Uygur stimuli	Russian script	Stress location in Russian	Stress perceived by Uyghur	gloss
BAza	'база	initial	initial	‘base’
DAcha	'дача	initial	initial	‘villa’
chaNA	'сани	initial	final	‘sledge’
BANkir	бан'кир	final	initial	‘banker’
banTIK	'бантик	initial	final	‘tie’
PARnik	пар'ник	final	initial	‘greenhouse’
KASsir	кас'сир	final	initial	‘accountant’
kasTUM	кос'тюм	final	final	‘suit’
PUTbol	фут'бол	final	initial	‘soccer’

Only the words that had etymology source are listed

In Table 21, for example, *dacha*, ‘villa,’ originally from Russian 'дача (initial stress), has first syllable stress. In Uyghur, if *dacha* was assimilated (nativized) in Uyghur when it was borrowed from Russian, we would predict *daCHA*, instead of *DAcha,* due to the stress pattern in Uyghur under the assumption that stress is on the final position. In Uyghur, we had *DAcha*, which is consistently perceived as a first syllable stressed word which was similar to the stress pattern in Russian. Uygur stress assignment on *BAza* was similar with the Russian stress pattern. However, the word *chaNA*, which was probably from 'сани ‘sledge’ did not borrow the stress from Russian. It could be the reason there were vowel changes. Only looking at Russian in which the Uyghur had contact in recent history, it seems that Uyghur did perceive the duration in CV syllable structure as a cue to the original Russian stress (the word *chaNA* ‘sledge’ 'сани was exceptional, but there is also a vowel change); however, this was not the case in CVC syllable structure, because the original stress and perceived stress pattern by Uyghurs were not the same in the CVC syllable structure. It could be possible that Uyghurs perceived the Russian stress and assigned stress based on duration only in the CV syllable structure.

There were no secondary long vowels in Uyghur. The duration effect was due to purely stress effects rather than secondary vowel length effects.

(3) Persian borrowings in our experiments

Persian loans used in our experiments are listed in Table 22 below (full citation see Appendix B). In terms of Persian borrowed words, Persian does not have a vowel length distinction: the short/long distinction in Modern Persian (/a, i, u/ are long and /e, æ, o/ are short), shows that vowel length is not phonemic in Modern Persian (Mahootian, 1997:309). In Modern Persian, lexical stress for at least nouns, compound nouns and adjectives falls word-finally (Ferguson, 1957; Kahnemuyipour, 2003).

Table 22 Persian /Tajik Borrowed Words with Modern Uyghur Variety

Modern Uyghur	Gloss in Modern Uyghur	Origin	Modern Persian/Tajik form	Modern Persian/Tajik gloss
PAchaq	leg	Persian	paacheh	'leg'; (Gilani, 1999, pp.253)
baHA	price	Persian	bahaa	'price' .(Gilani, 1999, pp.273)
DAka	gauze	Persian /Tajik	дока(doka)	'gauze' (Mamatnazarov, 2011, pp.360)
CHAsa	square	Persian	chahar suu	'square' (Steingass, 1973x, pp. 404)
DOra	medicine	Persian /Tajik	daaru дору (doru)	'medicine' (Gilani, 1999, pp.279) Tajik source (Mamatnazarov, 2011, pp.515)
SOda	business	Persian	soda saudaa	'bargain' (Ayranpur-Kashani and Ayranpur-Kashani, 1986, pp.18); *saudaa* 'marketing'(Steingass, 1973x, pp. 707)

Only the words that had etymology source are listed

Since Modern Persian has word-final stress, Uyghur participants appear to have assigned different stress location. For example, *DAka* 'gauze' and *DOra* 'medicine' are cognate to Modern Persian (or Tajik) *doKA* and *daaRU* (*daaru* had some phonetic changes). Uyghur has initial stress perceptually, while these words in Modern Persian have final stress. We don't know much about stress assignment in pre-modern Persian-when these words were borrowed into the antecedent language of Modern Uyghur. Modern Persian studies about stress began about 1933 (Windfuhr, 1979). In addition, compound words such as *chaar suu* 'square' had stress on the final positions in Persian. However, in Uyghur, these words are stressed on the first syllables regardless of original stress patterns. Even though long vowels in Persian did not attract stress, the Persian long vowels borrowed into Uyghur are stressed. For example, the Persian word *paacheh,* 'leg' which has a long vowel in the first syllable, borrowed into Uyghur as *PAchaq* 'leg' and the first syllable is stressed, but in Persian, stress was on the final syllable. The longer duration of the first vowel in *paacheh* 'leg' than the short vowel in the second syllable in that word could provide the stress cues in Uyghur. If Persian long vowels coincidently occur in final position, these long vowels hold the stress position both in the Persian origin and the Uyghur borrowing. Therefore, long vowels in Persian were perceived as stressed in Uyghur, but not the Persian origin per se if it occurred in non-final position.

(4) Arabic loans

There were three words from Arabic in Experiment 2. These disyllabic nouns contrasted in Uyghur perceived stress assignments as shown in Table 23 below. In Arabic, stress is assigned in several ways: (1) stress is assigned to the final syllable if it is superheavy (CVVC or CVCC), otherwise stress never falls on the final position; (2) stress falls on the penult if it is heavy (CVC or CVV); (3) usually stress falls on the antepenult position, not beyond the antepenult position (McCarthy, 1979, Fisher, 2001, p.20).

Table 23 Arabic borrowed words with Modern Uyghur (Wehr 1971)

Modern Uyghur	Gloss in Modern Uyghur	Modern standard Arabic from	Modern Standard Arabic gloss
baLA	disaster	bałaaʔ	‘disaster; misfortune’
jaPA	hardness	jafaaʔ	‘roughness, hardness; antipathy’
JAza	punishment	jazaaʔ	‘penalty, punishment’
shiPA	cure	shafaaʔ	‘cure’

In the Arabic loan, *jaPA*, ‘hardness’, stress in Modern Uyghur falls on the second syllable, which is similar to Modern Uyghur, corresponding to the phonemic long vowel in Arabic. However, in the Arabic loan *JAza* ‘penalty, punishment’, stress in Modern Uyghur falls on the first syllable rather than the second as syllable even though the second syllable in Arabic has a phonemic long vowel. In other words, Uyghur participants’ assigned stresses on *JAza* were different from the original one in Arabic. Both stimuli were from the CV-inconsistent group, which means participants had variability in assigning the stress. The CV-inconsistent group was chosen because of inconsistency of the perceptual stress evaluations by Native Uyghur speakers. The inconsistency is dependent on the first syllable of the disyllabic nouns that contrasted in the first syllable in terms of stress location; the second syllables were usually very consistent. In other words, when we indicated the inconsistency, it is based on the first syllables. The Arabic loan *JAza* is inconsistently indicated as first syllable stressed which means that Uyghur native speakers could assign the stress on the second syllable as *jaZA*. Assigning stress in Arabic loans is similar to Persian loans, the long vowels were perceived as stressed in Uyghur. It indicated the duration is a cue for both Arabic and Persian loans.

The Arabic *baLA* and *shiPA* likely came to eastern Turkic (including Uyghur) from Arabic through Persian (Dwyer, p.c). In Persian, the word ‘disaster’ is *bala:* and ‘cure’ is *shafaa:* both with had final long vowels. Stress of these two words also was in the final position in

Persian; therefore, these two words are stressed in final position in Uyghur. In addition, if these words were borrowed from Arabic, Uyghur kept the stress position the same as Arabic because these words had superheavy final syllables and stress was in the final position. *Bala* in Experiment 1 did not influence the results: after removing the *baLA-BAla* pair and running the statistics again, we found that duration and intensity provided cues for stress location. The Arabic loan *shafaa* in Experiment 2 did not influence the result because we focused on the first syllables, *sha*, of *shafaa*. In Experiment 2, potential vowel lengthening did not influence the results.

For the Russian, Persian and Arabic loans, Uyghurs assigned stress on heavy syllables or kept it the same as the original source. There are some other Russian loans in which Uyghur assigned stress on the heavy syllables regardless of original stress. The word borrowed from Russian ав'тобус ‘bus’ has the stress on the second syllable in Russian, but *aptobus* ‘bus’ in Uyghur has the stress on the last syllable which is heavy. However, when CV'CVC or 'CVC.CV structured words from Russian are borrowed in Uyghur, the original stress is maintained as shown examples as follow:

Table 24 Russian borrowed words with Modern Uyghur

Modern Uyghur	gloss	Russian	transliterate
béLET	‘ticket’	би'лет	bilet
piKAP	‘car’	пи'кап	pikap
gaRAZH	‘garage’	га'раж	garaȥ
magiZIN	‘shop’	мага'зин	magazin
pemiDOR	‘tomato’	поми'дор	pomidor
PARta	‘desk’	'парта	parta
boGHALtir	‘accountant’	бух'галтер	buxalter
poPAYka	‘sweater’	фу'файка	fufajka

http://feb-web.ru/feb/mas/mas-abc/default.asp 5th of March 2013

From these examples, it seems that Uyghurs indicate stress on heavy syllables for Russian words. Uyghurs accepted the stress of loan words as either the original one or stress on heavy syllables.

In conclusion, in words from Russian, Uyghur kept the original stress pattern in CV structure but not in CVC structure. In words from Persian, Uyghurs assigned stress on the long vowels rather than borrowing stress from the original Persian stress patterns. For Arabic borrowed words, Uyghurs kept the long vowels or heavy syllables stressed. The small sample, and inconsistent results do not allow a strong conclusion, but stress borrowing should be investigated in future research.

6.1.2 The Acquisition of Stress Patterns in Uyghur by Non-Native Speakers

We examined the acquisition of the stress pattern in Uyghur by non-native speakers to determine which acoustic parameters are used by Uyghur learners in producing Uyghur words contrasting in stress. The underlying question was whether non-native speakers learn the L2 stress pattern or transfer the L1 stress cues into the L2. The fundamental frequency, duration and intensity were measured in vowels from the stressed and the unstressed conditions. From the previous research, stress cues can transfer from L1 to L2. In addition, high proficiency speakers may learn native-like stress-patterns regardless of the similarity and dissimilarity of the L1 and the L2 (Zhang et al, 2008; Zuraiq & Sereno, 2007). Based on the selected stress cues by non-native learners, we can determine whether the stress pattern of learners results from the transfer from the L1 to the L2 or the independent learning process of the L2 stress pattern.

In this research, we examined minimal pairs (Experiment 4a) and disyllabic nouns contrasting stress on the first syllable (Experiment4b) in the production of Uyghur learners. From

Experiment 4a, the pattern of stress in Uyghur was neither like English nor Uyghur for the non-native speakers. These learners used a rising F0 for initial syllables regardless of stress location, which is more similar to English stress cue; therefore, we could say that they transfer F0 from the L1 into the L2. However, they applied the duration cue on the final (second) syllable regardless of stress location, placing stress in Uyghur in the final position as they were taught in the classroom. This shows that a rule-based stress pattern can be easily acquired. In terms of intensity, the learners did not use intensity, which was not similar to Uyghur or English patterns in producing minimal pairs.

In Experiment 4b, we used relatively higher frequency disyllabic nouns in Uyghur that were contrasted in terms of stress location on the first syllable. The pattern of the Uyghur learners was similar to Uyghur native speakers. Non-native speakers used duration as a cue for stress location, and did not use F0 and intensity as cues that were required in English stress pattern. These results suggest that stress pattern can be acquired by non-native speakers. Even if the stress cues in L1 and L2 were not similar, highly proficient speakers could manage to use native-like cues to produce native-like stress pattern.

In terms of L1 transfer to L2, if the transfer hypothesis holds, the following assumptions are possible:

(1) Native English speakers use F0, duration and intensity in producing Uyghur minimal pairs or disyllabic nouns that contrast in the first syllable.
(2) Lexical class does not matter in Uyghur, but it matters in English. For disyllabic nouns, native English speakers who learned Uyghur as an L2 should have initial stress

preferences regardless of stress location in Uyghur. In English, 76% of disyllabic nouns have initial stress (Sereno, 1986).

(3) English has complex syllable structures, and stress is related to the syllable structure (Guion et al, 2004). If the transfer assumption is correct, English learners of Uyghur should stress CVC syllable structures more often than CV syllable structures.

The results from Experiment 4a and 4b did not support any of the transfer assumptions. English learners of Uyghur did not use F0, duration and intensity to distinguish the stressed syllables from unstressed syllables. For low frequency words, which were the minimal pairs in Experiment 4a, English learners of Uyghur showed an intermediate pattern between L1 English and L2 Uyghur in stress pattern. They used a higher F0 and intensity in initial syllables regardless of stress location in Uyghur. They treated these words as disyllabic English nouns that have initial syllable stress. However, they used duration (the result of final lengthening or rule-based teaching) in the final position of the word. The present data cannot distinguish between these. In Experiment 4b, English learners of Uyghur used duration similar to native Uyghur speakers, and they did not use F0 and intensity. These results indicated that English learners of Uyghur did not transfer lexical class and acoustic parameters from English. In contrast, they successfully learned Uyghur stress patterns in which duration provides a systematic stress cue.

Our tentative explanation for the slightly different pattern of data in Experiment 4a and 4b is that the stimuli in Experiment 1 4a (and Experiment 1) were less common in Uyghur. In Experiment 4b, relatively higher frequency words were used (see appendix C). It could be the case that for lower frequency words people tend to be influenced from their native language. In other words, the lack of familiarity with the stimuli could be a reason. For high frequency words, the results may be similar to native speakers; only lower frequency words cause different results as compared to native speakers.

The differences across Experiment 4a and 4b are also not consistent with rule-based learning effects. In Uyghur, the regular stress is located in the final position of a word (Johanson, 1998). If English learners of Uyghur learned the stress rule in Uyghur, they should apply duration in both experiments. However, in Experiment 4a and 4b, the learners behaved differently. Therefore, the results cannot be due to rule-based learning. From the results of Experiment 4b, we tentatively suggest that stress may be stored in the lexicon for non-native speakers.

6.2 Conclusion

The present series of experiments provided an acoustic analysis of the stress pattern in Uyghur. They include minimal pairs and disyllabic nouns, the interaction of lexical stress with sentential intonation, as well as an examination of non-native speakers learning Uyghur. Duration is a strong cue in distinguishing the stressed syllables from the unstressed syllables. Uyghur does not use F0 and the role of intensity was less pronounced. Uyghur does not use pitch for distinguishing stress accent. This research suggests that Uyghur is not a pitch-accent language. Moreover, non-native speakers of Uyghur can learn the stress pattern of Uyghur. They produced relatively high frequency words that contrasted in stress location similar to native speakers. In other words, they also used duration as a stress cue, but they did not use F0 and intensity as stress cues.

6.3 Limitations

While the present acoustic research may provide a baseline for a phonological stress analysis of Uyghur, there are a number of limitations. The limitations are primarily due to the fact that access to Uyghur speakers and learners was difficult and that there is very little research on the phonetics and phonology of Uyghur. We only focused on a small number of disyllabic nouns. The sample size of stimuli could be larger and we did not investigate verbs. In addition, this research had a small number of participants both for Uyghur native speakers and Uyghur learners, allowing for subject variability. Moreover, for the English learners of Uyghur, proficiency level was not tested because there is no standard proficiency level test of Uyghur. Last but not least, perception of stress by native and non-native speakers was not examined. These perception studies could shed light on the cues native and non-native listeners use.

6.4 Pedagogical Implications

The present experiments may provide some insight into the teaching of Uyghur. Much research has focused on the production and perception of stress patterns by non-native speakers; however, little research directly studied how L2 teachers teach L2 accent patterns. Even though English has a body of materials for teaching grammar, far less material exist for teaching stress patterns (Derwing & Munro, 2005). When L2 learners produce the L2 words, they not only have difficulty in producing some phonemes that do not exist in their L1, but they also have difficulty in producing stress patterns in the L2, especially when the L1 does not use the same acoustic parameters.

L2 speakers commonly misplace the stress in their L2 (Archibald, 1992, Mennen, 2006). Munro and Derwing (1995) indicated that misplacing stress in the utterances may cause several issues. The lack of comprehensibility shows that the L2 utterance was hard to understand for L1 speakers. The lack of intelligibility (transcribe-ability) indicates which extended L2 utterance was understandable. The presence of a foreign accent means that the speech sample sounds like foreigners' speech, being different from native speakers. In order to tease apart these three dimensions, Munro and Derwing investigated Cantonese, Spanish, Polish and Japanese intermediate to high level ESL students who narrated a story. The authors selected 19 critical words from their speech. Native English speakers evaluated the words in terms of comprehensibility, intelligibility and accentedness. They found that native speakers of English were harsher on foreign accent rating than on comprehensibility and they were least harsh on intelligibility. For native English speakers, intelligibility is the key to understand, but native English speakers are stricter on accents than on comprehensibility and intelligibility. This indicated that accent rating was different from intelligibility. The research clearly indicated that these three dimensions were partly independent. Based on these results, Munro and Derwing suggest that intelligibility is a realistic goal to achieve in L2 teaching as a first step. The prosodic information about accent, pitch, and stress is required at a superior level (among the levels of novice, intermediate, advanced and superior) by ACTFL (American Council on Teaching of Foreign Language). At the superior level, L2 speakers should produce native -like stress, and sentential accent as shown below:

"Superior speakers command a variety of interactive and discourse strategies, such as turn-taking and separating main ideas from supporting information through the use of syntactic and lexical devices, as well as intonational features such as pitch, stress and tone." (ACTFL)

However, the acquisition of stress and intonation is a process and should be included from the beginning as a practical goal. In order to increase intelligibility, there are certain methods that may be possible to improve L2 speakers' production.

First, correcting them immediately indicates that when we are teaching L2, we should correct in the moment and let students repeat the correct form of mispronounced or misplaced stress accents, where students have mispronunciations. For example, when Uyghur learners produce misplaced stress accent at the word level, we provide the correct version immediately. In teaching, teachers should exaggerate the stressed syllable in terms of duration and let students focus on stress patterns in Uyghur. It may be helpful to acquire the correct form immediately. From my personal experience of learning English, I would like native speakers to correct me, highlighting the appropriate use of cues when I mispronounce words, and then I would pay more attention next time.

Secondly, we could also use a combination of perception and production training. In the teaching, we provide a situation in which L2 learners perceive the correct stress pattern and produce it later. When we create practice materials, we should focus on duration and ignore F0 and intensity, which are cues in Uyghur. L2 learners will learn which cues they should use when they produce the correct forms. Herd (2011) investigated the acquisition of Spanish phonemes by English speakers. In her research, she used three different training methods including only production, only perception, and a combination of production and perception. Herd (2011) found that L2 learners had more benefits from only perception and a combination of production and perception training. She concluded that perception training improves their production in L2. Suprasegmental learning seems to be similar to segmental learning in L2 (Trofimovich & Baker, 2006). The L2 training effect was observed not only at the segmental level but also at the suprasegmental level. Wang et al (1999) trained English learners of Chinese to acquire Mandarin Chinese tone. They examined the acquisition of tone in Mandarin Chinese by non-native speakers. They included pre-test, training, immediate post-test and delayed post-test (after six months). They found that participants acquire the tones to some extent (shown also as changes at the neural level) and that even after six months, trainees still kept the Mandarin tone production improvements. Wang et al. (2003) also found that perception based training can substantially improve the production by 18 % in Chinese tone perceptual training. It could be possible to use the perception and production training in the acquisition of Uyghur stress pattern for non-native speakers. Another merit of this method is that L2 learners will learn stress patterns actively and they will learn from different speakers and from themselves. Sometimes, the first option, 'immediate correcting,' does not work for some shy L2 leaners, but the second method may be effective for this type of L2 learner.

The perception and production training could use a strict lab-based practice using discrimination tasks and identification tasks in which teachers provide minimal pairs and ask students to indicate where the stress is, or students could listen to word pairs and indicate whether the word pair has the same stress or different stress. In order to make tasks close to natural speech, we can use multiple speakers. In order to use appropriate acoustic parameters, L2 Uyghur teachers could slightly modify the acoustic parameters as well as the exposure of the non-native speakers to the stimuli. Uyghur teachers could encourage L2 learners to use the correct acoustic parameters for Uyghur. For example, English uses fundamental frequency, duration and intensity; however, Uyghur uses only duration. In Uyghur teaching, the purpose of the tasks is to let the students focus on duration and ignore F0 and intensity. Such laboratory methods may aid learners in acquiring native-like accent patterns in Uyghur.

References

ACTFL (American Council on the Teaching Foreign Languages): Proficiency Guideline-Speakin, Retrivied from http://www.languagetesting.com/scale.htm (Feb8th 2013)

Altmann, H. (2006). *Perception and Production of Second Language Stress*.(Doctoral dissertation), Retrieved from ProQuest Digital Dissertations. (UMI No: AAT 3220627)

Archibald, J. (1992). Transfer of L1 Parameter Settings: Some Empirical Evidence from Polish Metrics. *Canadian Journal of Linguistics*, 37, 301–339.

Archibald, J. (1993). The Learnability of English Metrical Parameters by Adult Spanish Speakers. *International Review of Applied Linguistics and Language Teaching*, 31/32, 129-142.

Archibald, J. (1997). The Acquisition of English Stress by Speakers of Non-accentual Languages: Lexical Storage versus Computation of Stress. *Linguistics*, 35, 167-181.

Archibald, J. (1998). Second Language Phonology, Phonteics and Typology. *Studies of Second Language Acquistion*, 20, 189-211.

Aryanpur-Kashani, A., & Aryanpur-Kashani, M. (1986). *The Combined New Persian-English and English-Persian Dictionary*. Mazda Publishers.

Beckman, M. (1986). *Stress and Non-stress Accent*: Foris Pub USA.

Beckman, M., & Pierrehumbert, J. (1986). Intonational Structure in Japanese and English. *Phonology Yearbook*, 3(1), 5-70.

Boersma, P., & Weenik, D. (2011). Praat, a system for doing phonetics by computer.Retrivied from www.praat.org.

Clauson, G. (1972). *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*. Clarendon Press.

Cutler, A. (1980). Errors of Stress and Intonation. In V.A. Fromkin (Ed.), *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand.* New York: Academic Press, 67-80.

Delattre, P. (1969): *The General Phonetic Characteristics of Languages. An Acoustic and Articulatory Study of Vowel Reduction in Four Languages*, Final Report, University of California, Santa Barbara, CA, USA.

Derwing, T. M., & Munro, M. J. (2005). Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, 39(3), 379-397.

Dobrovolsky, M. (1999). The Phonetics of Chuvash Stress: Implications for Phonology. *In Proceedings of the 14th International Congress of Phonetic Sciences*,Vol. (1), 539-542.

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.

Dwyer, A. M. (2000). Consonantalization and Obfuscation. In Goksel, Asli & Celia Kerslake (eds.) *Studies on Turkish and Turkic Languages*. Proceedings of the Ninth International Conference on Turkish Linguistics. Lincoln College, Oxford, 12-14 August 1998 (Turcologica 46). Wiesbaden: Harrassowitz, 423-432.

Dwyer, A. M. (2002). The Turkic Languages. In Levinson, David and Karen Christensen, eds. *Encyclopedia of Modern Asia*. New York: Scribner's.

Erdmann, P. H. (1973). Patterns of stress-transfer in English and German. *IRAL - International Review of Applied Linguistics in Language Teaching*. Volume 11, Issue 1-4, Pages 229–242, ISSN (Online) 1613-4141, ISSN (Print) 0019-042X, DOI: 10.1515/iral.1973.11.1-4.229, November 2009.

Fant, G., & Kruckenberg, A. (1994). Notes on Stress and Word Accent in Swedish. *Speech Transmission Laboratory Quarterly Status Report*, 35(2-3), 125-144.

Ferguson, C. A.(1957). Word Stress in Persian, *Language*, 33, 123-135.

Fisher, W., (2001). *A Grammar of Classical Arabic*. Translated by Rodgers, J., (2001); Yale University Press.

Flemming, E. (2005). A Phonetically-based Model of Phonological Vowel Reduction. Retrivied from: http://web. mit. edu/~ flemming/www/paper/vowelred. pdf.

Fourakis, M. (1991). Tempo, Stress, and Vowel Reduction in American English. *The Journal of the Acoustical Society of America*, 90(4), 1816-1827.

Fry, D. B.(1955). Duations and Intensity as Physical Correlates of Linguistic Stress. *The Journal of the Acoustical Society of America*, 27(4),765-768.

Fry, D. B. (1958). Experiments in the Perception of stress. *Language and Speech*, 1 (2), 126-152.

Gilani, D., (1999). *An English Persian Dictionary*. IBEX Publisher

Guion, S. G., Harada, T., & Clark, J. J.(2004). Early and Late Spanish-English Bilinguals' Acquisition of English Word Stress Patterns. *Bilingualism: Language and Cognition,* 7(3), 207-226.

Gordon, M. (2004). A Phonological and Phonetic Study of Word-Level Stress in Chickasaw. *International Journal of American Linguistics*, 70(1), 1-32.

Jaing, H., Liu, Y., & Lu, L. (2010). 维吾尔语词重音实验研究。《民族语文》，第三期，67-71.[An Experimental study of stress in Uyghur words. *Ethnic Study*, Vol (3)]

Johanson, L. & Csato É. Á. (1998). *The Turkic Languages*. Routledge.

Jun, S. A. (2005). Prosodic typology. In Jun, S.A (Ed.). *Prosodic Typology: the Phonology of Intonation and Phrasing* . Oxford: University Press, 430-458.

Hamilton, W. S. (1980). *Introduction to Russian Phonology and Word Structure*. Slavica Publishers.

Halle, M., & Idsardi, W. (1994). General Properties of Stress and Metrical Structure. *DIMACS Serier of Discrete Mathematics and Theoretical Computer Science*, 17, 37-70.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*: University of Chicago Press.

Herd, W. (2011). *The Perceptual and Production Training of /d, tap, r/ in L2 Spanish: Behavioral, Psycholinguistic, and Neurolinguistic Evidence.* (Doctoral dissertation), retrieved from ProQuest Digital Dissertations. (UMI No: AAT: 3457613)

Hirst, D., & Cristo, A. (1998). A Survey of Intonation System. *Intonation System: A Survey of Twenty Languages.* Edited by Hirst, D., and Cristo, A. Cambridge University Press, 1-44.

Hualde, J., Elordieta, G., Gaminde, I., & Smiljanic, R. (2002). From Pitch Accent to Stress Accent in Basque. *Laboratory Phonology*, 7, 547-584.

Huss, V., (1978). English Word Stress in the Post-Nuclear Position. *Phonetica*, 35, 86-105.

Hyman, M. L.(2006). Word Prosody Typology. *Phonology*, 23, 225–257.

Kahnemuyipuor, A. (2003). Syntactic Categories and Persian Stress. *Natural Language & Linguistic Theory*, 21(2), 333-379.

Katzner, K., (1994). *English-Russian, Russian-English Dictionary*. Wiley.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness Predicts Prominence: Fundamental Frequency Lends Little. *The Journal of the Acoustical Society of America*, 118(2), 1038–1054.

Kodzasov, S. V. (2003). Towards an Evolutionary Typology of Stress. In *Proceedings of the 15th International Congress of Phonetic Sciences*, 63-66.

Kuznetsova, E. V. (2006). Acoustic Parameters of Secondary Stress in Russian. In *Proceedings of the XVIII Meeting of the Russian Acoustical Society*, 547-550.

Ladd, D. R. (2008). *Intonational Phonology*. Cambridge University Press.

Lai, Y., (2008). *Acoustic Realization and Perception English Lexical Stress by Mandarin Learners.* (Doctoral dissertation), Retrieved from ProQuest Digital Dissertations. (UMI No: AAT 3316252)

Levi, S. (2005). Acoustic Correlates of Lexical Accent in Turkish. *Journal of the International Phonetic Association*, *35*(01), 73-97.

Liang.J.,and Zhang,J.(2008).维吾尔语双音节词重音的产出实验研究。《第八届中国语音学术会议暨庆贺吴宗济先生百岁华诞语言科学前沿问题国际研讨会论文集》[A Production Study on Stress Pattern on the Disyllabic Words in Uyghur Language. *8th Chinese Phonetics Conferences and Languge Science International Symposium on Fronter Issues in the honor of Celebrating Mr. Wu Zongxian's Centary Brithday.*]

Lindblom, B. (1963). Spectrographic Study of Vowel Reduction. *The Journal of the Acoustical Society of America*, 35, 1773-1781.

Lindström, E., & Remijsen, B. (2005). Aspects of Prosody of Kuot, a Language Where Intonation Ignores Stress. *Linguistics*, 43(4), 839-870.

Mahootian, S. (1997) *Persian*. Routledge.

Mamatnazarov, A., (2011). *Standard English-Tajik Dictionary*. Düshenbe.

McCarthy, J. J. (1979). *On stress and syllabification*. Linguistics Department Faculty Publication Series. http://scholarworks.umass.edu/linguist_faculty_pubs/53

Mennen, I. (2006). Phonetic and Phonological Influences in Non-native Intonation: an Overview for Language Teachers. *Phonetic and Phonological Influences in Non-native Intonation: an Overview for Language Teachers.* (WP-9).

Munro, M. J., & Derwing, T. M. (1995). Processing Time, Accent, and Comprehensibility in the Perception of Native and Foreign-Accented Speech. *Language and Speech*, 38(3), 289-306.

Muti, I. (2007). Uyghur Tilidiki Söz Urghusi we Ibare Intonatsiyisi---Memlikitimizdiki Turkiy Tillar we bu Heqtiki Ilmiy Tetiqiqat **[**Lexical Stress and Phrase Intonation in Uyghur – Scientific Research on Turkic Languages and the Research about this in Our Country**].** *<<Ibrahim Muti maqaliliri>>* [The articles of Ibrahim Muti]. Milletler Neshiryati, Beijing, 40-59.

Nash, C. (2005). A Phonetic and Phonological Approach to Stress in Sinhala Verbs.*In the Workshop on Sinhala Linguistics, University of California, Santa Barbara, June 4-5*. Retrieved from http://www.linguistics.ucsb.edu/research/cmnash_2005_stress_verb.pdf.

Ortega-Llebaria, M., & Prieto, P. (2011). Acoustic Correlates of Stress in Central Catalan and Castilian Spanish. *Language and Speech*, 54(1), 73-97.

Osmanov, M. (2009). Uyghur Tilidiki Uzun Sozuq Tawushlar Mesilisi Toghrisida. *Bulaq*, 129 (6), 56-78. [The Issues about the Long Vowels in Uyghur. Bulak]

Özçelik. Ö. (2012). *Representation and Acquisition of Stress: the Case of Turkish*. (Doctoral Dissertation), retrivied from http://digitool.Library.McGill.CA:80/R/-?func=dbin-jump-full&object_id=107766&silo_library=GEN01, PID: 107766

Pater, J. V. (1997). Metrical Parameter Missetting in Second Language Acquisition. In S. J. Hannahs & M. Young-Scholten (eds.), *Focus on Phonological Acquisition*. Amsterdam: John Benjamins, 235–261.

Saimaiti, M., & Feng, Z. (2007). A Syllabification Algorithm and Syllable Statistics of Written Uyghur. *In Proceedings of the 4th Corpus Linguistics Conference,* Birmingham, UK. Retrivied from http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2007/153Paper.pdf

Sereno, J. A. (1986). Stress Pattern Differentiation of Form Class in English. *The Journal of the Acoustical Society of America*, 79, S36.

Sereno, J.A. and Jongman, A. (1995). Acoustic Correlates of Grammatical Class. *Language and Speech*, 38, 57-76.

Sluijter, A. M. C., & Van Heuven, V. J. (1996 a). Spectral Balance as an Acoustic Correlate of Linguistic Stress. *The Journal of the Acoustical Society of America*, 100(4), 2471-2485.

Sluijter, A. M., & Van Heuven, V. J. (1996 b). Acoustic Correlates of Linguistic Stress and Accent in Dutch and American English. In *Spoken Language, 1996. ICSLP 96. Proceedings, Fourth International Conference on,* Vol(2), 630-633. Retrivied from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=607440.

Steingass, F., (1973x). *Persian-English Dictionary*. New Dehli-110055: Orient Book Reprint Corporation

Trofimovich, P. & Baker, W. (2006). Learning Second Language Suprasegmentals: Effects of L2 Experience on Prosody and Fluency Characteristics of L2 Speech. *Studies in Second Language Acquistion*, 28, 1-30.

Van Bergem, D. R. (1993). Acoustic Vowel Reduction as a Function of Sentence Accent, Word Stress, and Word Class. *Speech Communication*, 12(1), 1-23.

Van Bergem, D. R. (1995). Perceptual and Acoustic Aspects of Lexical Vowel Reduction, a Sound Change in Progress. *Speech Communication*, 16(4), 329-358.

Van Der Mark, S. C. (2003). *Phonetics in Blackfoot Pitch Accent.* M.A., The University of Calgary, Calgary. Retrivied from http://hdl.handle.net/1880/40174.

Wade, T. (1992). A Comprehensive Russian Grammar. Oxford, UK: Blackwell

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American Listeners to Perceive Mandarin Tones. *The Journal of the Acoustical Society of America*, 106(6), 3649-3658.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and Perceptual Evaluation of Mandarin Tone Productions before and after Perceptual Training. *The Journal of the Acoustical Society of America*, 113, 1033-1043.

Wehr, H. and Cowan, M. (1971). *A Dictionary of Modern Written Arabic*. Spoken Language Service, Itacha, NewYork

Xinjiang Weiwuer zizhiqu renmin zhengfu [Xinjiang Uyghur Autonomous Region People's Government] (eds.). 2008. Xinjiang nianjian [Xinjiang Yearbook]. Urumqi: Xinjiang Nianjian Publish House.

Yakup, A. and Gheyurani, Gh. (1999). *Uyghur Tilining Izahliq Lughiti*. [Uyghur Explanatory Dictionary]. Xinjiang People Press.

Zhang, Y., Nissen, S. L., & Francis, A. L. (2008). Acoustic Characteristics of English Lexical Stress Produced by Native Mandarin Speakers. *The Journal of the Acoustical Society of America*, 123(6), 4498-4513.

Zuraiq, W. and Sereno, J.A. (2007). English Lexical Stress Cues in Native English and Non-native Arabic Speakers. In J. Trouvain and B. Barry (Eds.) *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS XVI)*, 829-832.

Zuraiq,W. (2005). *The production of Lexical Stress by Native Speakers of Arabic and English and by Arab Learners of English.* (Doctoral dissertation), retrieved from ProQuest Digital Dissertations. (UMI No: AAT 3203258)

Appendix A

1. Questionnaire for Uyghur native speakers background information

Questionnaire

Gender: __________

Age: __________

Native country/state _______________

Native language ______________

Knowledge of OTHER languages: Write the name of the language in the blank, and indicate your approximate abilities in each of the four areas for each language.

1. Language: ________________________

Speaking	Listening	Reading	Writing
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

2. Language: ________________________

Speaking	Listening	Reading	Writing
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

3. Language: ________________________

Speaking	Listening	Reading	Writing
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good

□ Near-Native □ Near-Native □ Near-Native □ Near-Native

4. What was your age when you started learning English? _______________

5. Did you take English in Elementary School? Yes No

If yes, where? ___________________

For how many years? ___________________

6. Did you take English in High School? Yes No

If yes, where? ____________________

For how many years? ____________________

7. Did you study English at the college level? Yes No

If yes, where? ___________________

For how many years? _____________________

8. Have you lived in an English speaking country? Yes No

If yes, where? ________________

For how many years? _____________________

9. Have you had any informal, out of classroom, exposure to English? Yes No

If yes, please mark all exposure you have had.

--------- Music in English.

--------- English speaking relatives.

--------- English speaking friends.

--------- Vacation travel to English speaking country.

--------- English languages magazines/ newspapers.

--------- English speaking TV.

10. Do you have a foreign accent in English? Yes No

If yes, please rate the strength of your accent.

□ No Accent □ Slight Accent □ Moderate Accent □ Strong Accent

Thank you for your participation.

2. Questionnaire for English native speakers' background information

Questionnaire

Gender: __________

Age: __________

Native country/state _______________

Native language ______________

Knowledge of OTHER languages: Write the name of the language in the blank, and indicate your approximate abilities in each of the four areas for each language.

1. Language: ________________________

Speaking	Listening	Reading	Writing
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

2. Language: ________________________

Speaking	Listening	Reading	Writing
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

3. Language: ________________________

Speaking	Listening	Reading	Writing
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

4. What was your age when you started learning Uyghur? _______________

5. Did you take Uyghur in Elementary School? Yes No

If yes, where? ___________________

For how many years? ___________________

6. Did you take Uyghur in High School? Yes No

If yes, where? ____________________

For how many years? ____________________

7. Did you study Uyghur at the college level? Yes No

If yes, where? ___________________

For how many years? _____________________

8. Have you lived in an Uyghur speaking community? Yes No

If yes, where? ________________

For how many years? _____________________

9. Have you had any informal, out of classroom, exposure to Uyghur? Yes No

If yes, please mark all exposure you have had.

--------- Music in Uyghur.

--------- Uyghur speaking relatives.

--------- Uyghur speaking friends.

--------- Vacation travel to a Uyghur speaking country.

--------- Uyghur languages magazines/ newspapers.

--------- Uyghur TV.

10. Do you have a foreign accent in Uyghur? Yes No

If yes, please rate the strength of your accent.

□ No Accent □ Slight Accent □ Moderate Accent □ Strong Accent

Thank you for your participation.

Appendix B

Experiment 1 Stimuli Etymology Table

Stimuli	gloss	origin	Origin spelling	Origin explanation	Cited source
Acha	elder sister	Old Turkic	eche:	‘one's mother's younger sister; one's own elder sister'; Xak.: 'elder sister;' Chagh.: 'an elderly woman'	Clauson, 1972, pp 20
aCHA	branching				
Ara	fork				
aRA	between	Old Turkic	ara: (ʔa:ra)	‘between'; Xak.: ara: 'in the middle of things'; Chagh.: ara,'middle, center'	Clauson, 1972, pp196
BAla	child	Old Turkic	bala:	'young bird, nestling'; Xak.: 'a nestling' metaphor for 'a young of any predatory animal'; (balu: with bala:)' a helper for a man in his work' (Kash); Chagh.: ' a young of animal'	Clauson, 1972, pp332
baLA	disaster	Arabic	balaa' بلاء ‘: is *hemze* ء	distress; misfortune	Dictionary of Modern written Arabic, by Wehr, 1971, p75
CHAtaq	bad branch of tress	Possibly Persian	chahaar tarkeh	‘chahaar’ means 'four' and ‘tarkeh’ means 'twig' ; ‘chahaar’ can be 'chaar' from http://dsal.uchicago.edu/cgi-bin/philologic/contextualize.pl?p.2.steingass.164975 (feb16th 2013).	English-Persian Dictionary by Gilani, 1999, p168, p549(separate meaning were from this cite)
chaTAQ	problem				
PAchaq	leg	Persian	paacheh	‘paacheh’ as noun/verb, meaning with 'leg'	English-Persian Dictionary by Gilani, 1999, p168, p253
paCHAQ	piece				
TÖshük	kitchen	Old Turkic	n/a	Kash.'s meaning does not seem to survive, uyg. VIII. (cf.Tütün) but exact meaning is unknown; (My guess) It is possible meaning of kitchen if it comes with tütün (smoke)	
töSHÜK	hole	Old Turkic	teshük /teshik	‘hole'; from 'tesh-; -lit. 'pierced'. Xak.: reptured'; 'a glutton'; Chagh.: 'hole'	Clauson, 1972, pp563

Experiment 2 Stimuli Etymology Table

stimuli	gloss	origin	Origin spelling	Origin explanation	Origin source Stress location	Cited source
BAza	base	Russian	'база	base; foundation; depot; camp	First syllable	English-Russian Russian-English dictionary by Kenneth Katzner, 1994,pp543
baHA	price	Persian	bahaa	price		English-Persian Dictionary by Gilani, 1999, p373,
DAka	gauze	Tajik	дока	gauze		Standard English-Tajik dictionary, by Mamatnazarov,2011, pp360
daLA	plain					
DAcha	villa	Russian	'дача	country house; summer cottage; dacha	First syllable	English-Russian Russian-English dictionary by Kenneth Katzner, 1995, pp618
daDA	father	Old Turkic	dede:	Oghuz XI.' father'; survive only in SW Osm.		Clauson, 1972, pp451
CHAsa	square	Persian	Chaar suu	square		from:http://dsal.uchicago.edu/cgi-bin/philologic/contextualize.pl?p.2.steingass.160486, feb 16th 2013
chaNA	sledge	Russian	'сани	sleigh; sled	First syllable	English-Russian Russian-English dictionary by Kenneth Katzner, 1995, pp958
DOra	medicine	Persian	daaru Tajik [doru]	medicine		English-Persian Dictionary by Gilani, 1999, p168, p279; Tajik: Tajik practical dictionary by Jilani, 2009
doQA	forehead					
POchi	boaster					
poTA	waist belt					
BANkir	banker	Russian	бан'кир	banker	Second syllable	English-Russian Russian-English dictionary by Kenneth Katzner, 1994,pp544
banTIK	bow	Russian	'бантик	small bow	First syllable	English-Russian Russian-English dictionary by Kenneth Katzner, 1994,pp544
KASsir	accountant	Russian	кас'сир	cashier; (bank) teller; ticket seller	Second syllable	English-Russian Russian-English dictionary by Kenneth Katzner, 1994,pp691
kasTUM	suit	Russian	кос'тюм	suit; outfit, attire; costume	Second syllable	English-Russian Russian-English dictionary by Kenneth Katzner, 1994,pp706
PARnik	greenhouse	Russian	пар'ник	hotbed		English-Russian Russian-English dictionary by Kenneth Katzner, 1994,pp829
parLASH	choosing					
SEPkün	mole					
sepDASH	fellow					

stimuli	gloss	origin	Origin spelling	Origin explanation	Origin source Stress location	Cited source
PUTbol	football	Russian	фут'бол	soccer	Second syllable	English-Russian Russian-English dictionary by Kenneth Katzner, 1994,pp1052
putLASH	trapping	Old Turkic	bu:t (bu:d)	‘the thigh’ sometimes more generally ‘leg’; Xak.: bu:t, ‘the thigh’; Chagh.: but ‘the leg from thigh to the toes’; *-la* is suffix functioning of verb; *-Ish* is gerund		Clauson, 1972, pp297
TOQmaq	stick	Old Turkic	tokı:mak	from 'tokı:-' meaning with 'hit, knock'; 'club, mallet' Xak.: 'mallet'; Chagh.: changed into 'tokmak' meaning ' well-known implement used to drive in tent pegs'		Clauson, 1972, pp471
toqQUZ	nine	Old Turkic	tokku:z	‘nine'; Xak.: toku:z; Chagh.: tokuz		Clauson, 1972, pp474
SOda	business	Persian	soda	bargain		The combined Persian-English and English-Persian dictionary, by Ayranpur-Kashani and Ayranpur-Kashani, 1986, pp18
soQA	plough	Old Turkic	soku:(sokghu:)	from 'sok-' meaning with 'beat, crush'; Xak.: soku: 'a mortar'; Chagh.: sokku: 'a large wooden mortar'		Clauson, 1972, pp805
KIno	movie	Russian	ки'но	motion picture; movies; cinema	Second syllable	English-Russian Russian-English dictionary by Kenneth Katzner, 1994,pp693

stimuli	gloss	origin	Origin spelling	Origin explanation	Origin source Stress location	Cited source
kiSHI	person	Old Turkic	kishi:	‘man, person, humanbeing'; Xak.: 'a man, mankind'; Chagh.: 'man' in geneal, female/male; 'a man' in singular		Clauson, 1972, pp753
JAza	punishment	Arabic	jafaa' جفاء	roughness, hardness; antipathy	Second syllable	Dictionary of Modern Written Arabic, by Wehr, 1971, p128
jaPA	hardness	Arabic	jazaa' جزاء	penalty, punishment	Second syllable	Dictionary of Modern Written Arabic, by Wehr, 1971, p124
SHIre	desk					
shiPA	cure	Persian	shafaa	‘cure’		English-Persian Dictionary by Gilani, 1999, p100,
CHOla	time					
choKA	chopstick					
TOxu	hen					
toPA	dirt, soil	Old Turkic	possibility1: to:pü:(to:po:)	the top ' of natural feature(mountain), hence 'hill'; variety: NW 'töbe' to SC Uzb. Tepa; Xak.: 'summit of a mountain'; Chagh. töpe 'a tall hill', 'the top of the head'		Clauson, 1972, pp436
toPA	dirt, soil	Old Turkic	Possibility 2: tüpi:	high wind' later high wind carrying snow or dust' Xak.: 'a high wind'		Clauson, 1972, pp436

The empty items indicated that their etymology sources are not available.

Appendix C

Frequency Table of Stimuli in Experiment 2

CV-consistent			
intial stressed	frequency	final stressed	frequency
DAka	328	daLA	2012
CHAsa	823	chaNA	146
BAza	7160	baHA	15752
DAcha	196	daDA	2325
DOra	22519	doQA	70
POchi	55	poTA	272
AVG	**5180**		**3430**
CVC-consistent			
intial stressed	frequency	final stressed	frequency
SEPkun	67	sepDASH	71
PUTbol	6302	putLASH	3
KASsir	221	kasTUM	159
PARnik	1083	parLASH	11
BANkir	54	banTIK	12
TOQmaq	1064	toqQUZ	8110
AVG	**1465**		**1394**
CV-inconsistent			
intial stressed	frequency	final stressed	frequency
SOda	36951	soQA	99
KIno	4715	kiSHI	31489
JAza	10314	jaPA	3001
SHIre	339	shiPA	1040
CHOla	19	choKA	446
TOxu	2585	toPA	5713
AVG	**9154**		**6965**