
Acoustic correlates of lexical stress in Uyghur

Mahire Yakup and Joan A. Sereno

Journal of the International Phonetic Association / *FirstView* Article / March 2016, pp 1 - 17
DOI: 10.1017/S0025100315000183, Published online: 20 January 2016

Acoustic correlates of lexical stress in Uyghur

Mahire Yakup
Nazarbayev University & University of Kansas
yakefu.mayila@nu.edu.kz

Joan A. Sereno
University of Kansas
sereno@ku.edu

The present study examined lexical stress patterns in Uyghur, a Turkic language. The main goal of this research was to isolate and determine which acoustic parameters provide cues to stress in Uyghur. A number of studies have investigated the phonetic correlates of lexical stress across the world's languages, with stressed syllables often longer in duration, higher in pitch, and greater in amplitude. The present study systematically investigated the acoustic cues to stress in Uyghur, examining duration, fundamental frequency, and amplitude. Three experiments were conducted: one utilizing minimal pairs in Uyghur, one examining disyllabic nouns in Uyghur that contrasted in the first syllable, and one investigating the interaction of lexical stress with Uyghur sentence intonation. The data consistently show that duration was a robust cue to stress in Uyghur, with less consistent effects for intensity. The data also clearly show that fundamental frequency was not a cue to lexical stress in Uyghur. Uyghur does not use the fundamental frequency to distinguish stressed from unstressed syllables. The results suggest that Uyghur does not pattern like a pitch-accent language (e.g. Turkish), but rather like a stress-accent language.

1 Introduction

Lexical stress patterns in languages have been instrumentally investigated for more than half a century (e.g. Fry 1955, 1958; Lieberman 1960; Beckman 1986). The main focus of these investigations has been to identify which acoustic parameters provide cues to assigning stress. Stressed syllables are often associated with one or more acoustic cues including duration, intensity, fundamental frequency (f0), and vowel quality. Not surprisingly, these acoustic parameters vary from language to language. While there is a relatively large literature examining the acoustic correlates of stress in English, there are very few quantitative acoustic studies investigating stress cues in non-Indo-European languages. The present study furthers the phonetic investigation of lexical stress by examining the acoustic cues to stress in Uyghur, an understudied Turkic language.

One of the first to acoustically analyze English stress was Fry (1955, 1958). Fry examined five minimal pairs (such as *SUBject* and *subJECT*) from twelve English native speakers. Since the 1950s, a number of studies have continued investigating the acoustic correlates of lexical stress in a variety of languages including English, Polish, French, Spanish, and Swedish

doi:10.1017/S0025100315000183

(Lieberman 1960, Lehiste 1970, Beckman 1986, Sereno 1986, Beckman & Edwards 1994, Turk & Sawusch 1996, Recasens & Espinosa 2006, Lai 2008, Zhang, Nissen & Francis 2008, Ortega-Llebaria & Prieto 2011). These studies have concentrated on four acoustic measures of perceived stress: duration, intensity, fundamental frequency, and spectral composition. In general, longer duration, greater amplitude, higher fundamental frequency, and less vowel reduction in a syllable have been shown to contribute to stress. Interestingly, Sereno & Jongman (1995) observed similar systematic acoustic differences, mainly in duration and amplitude, even in words (such as *answer*) that do not change their stress pattern with a change in grammatical class from a noun to a verb.

However, the individual contribution of each of these distinct acoustic parameters in signaling lexical stress varies across languages. Examining stress patterns in Dutch and American English, Sluijter & van Heuven (1996b) found that overall duration was the strongest cue to stress in both languages while measures of overall intensity as well as vowel quality were poor cues. They concluded that while overall intensity was not a consistent cue to stress, spectral balance (i.e. intensity distribution across frequency) systematically signaled stressed syllables, by means of intensity cues in the higher frequencies. In languages such as Spanish and Catalan (Ortega-Llebaria & Prieto 2011), duration cues were also very prominent as correlates of stress, while pitch, intensity and vowel quality contributed little. In polysyllabic Sinhala verbs, Nash (2005) found that combined duration and intensity information (syllable duration combined with intensity) distinguished stressed syllables from unstressed syllables. When syllable durations were equal, syllables that had a greater intensity were perceived as stressed; on the other hand, in an equal intensity situation, syllables that had longer duration were perceived as stressed. Both duration and intensity seem to contribute to stress in Sinhala.

While variations in duration, amplitude and formant structure systematically contribute to stress judgments, some studies (see Kochanski et al. 2005) find that fundamental frequency also contributes to perceived stress. However, the role of fundamental frequency as an acoustic correlate of stress seems to be intertwined with sentence-level accents (Sluijter & van Heuven 1996a, b), and in the absence of accents (fundamental frequency at the sentence level), fundamental frequency may lose its role as a cue to lexical stress.

It is not surprising that the relative effectiveness of fundamental frequency depends on the position of the word in the sentence. The importance of each of these acoustic parameters also varies with the position of the lexical item in the sentence (Morton & Jassem 1965, Gay 1978, Nakatani & Aston 1978, Beckman 1986). Early research by Huss (1978) measured the acoustic parameters of disyllabic minimal pairs in the nuclear position and post-nuclear position in English. Huss found that there was a difference between stressed syllables and unstressed syllables in terms of fundamental frequency in the nuclear position, with duration and amplitude as secondary cues; however, this difference disappeared in the non-nuclear position, though participants were still able to identify a word as a noun or verb. Moreover, in non-nuclear position, amplitude and duration cues became more salient, often surpassing fundamental frequency as a cue.

The inconsistent role of pitch can also be observed in languages other than English. Gordon (2004) tested eight native speakers in the production of polysyllabic words in Chickasaw, a Native American language. Comparing three stress levels using fundamental frequency, duration and intensity, he investigated whether the primary and secondary stressed syllables were acoustically different from each other and from unstressed syllables and found that words with short vowels did not show significant effects in fundamental frequency. He did find, however, that fundamental frequency was significant in distinguishing the primary stressed syllables from the secondary and the unstressed syllables in syllables only with long vowels. Moreover, intensity systematically differentiated primary stress from both secondary and unstressed syllables for both long and short vowels.

Interestingly, for other languages, studies have shown that there is no fundamental frequency role at all for distinguishing the stressed syllables from the unstressed syllables. Take the Kuot language (spoken in Papua New Guinea), for example. Measuring fundamental

frequency, duration and vowel quality in the first two syllables of a word, Lindström & Remijsen (2005) examined the interaction of the fundamental frequency contour with lexical stress in Kuot and found that, even when words were in different positions in the sentence, duration and vowel quality, but not fundamental frequency, provided cues to stressed syllables as compared to unstressed syllables. Non-significant fundamental frequency results showed that even though there were robust intonation changes, fundamental frequency was not as involved in distinguishing the stress location as other parameters (duration and vowel quality). The researchers concluded that Kuot is a language that uses duration and vowel quality to indicate lexical stress.

In languages like Japanese, however, stress is mainly realized by fundamental frequency variation (Beckman 1986, Flege & Hillenbrand 1984, Juffs1990, Chen et al. 2001, Zhang & Francis 2010) with very little information about stress provided by duration or intensity.

Beckman (1986) further examined the acoustic correlates of stress, elaborating on the contrast between stress-accent languages and pitch-accent languages (but see Hyman 2009). She showed that languages can use either multi-dimensional cues or a one-dimensional cue to differentiate stressed from unstressed syllables. Languages differ in the degree to which each acoustic cue contributes to the phonetic realization of stress. For Beckman, stress-accent languages, such as English, use duration, intensity and pitch information to distinguish stressed syllables from non-stressed syllables while pitch-accent languages, such as Japanese, primarily use one dimension – fundamental frequency – as the primary cue to distinguish stressed syllables from unstressed syllables.

In general, in stress languages, duration appears to be a robust cue for distinguishing stressed syllables from unstressed syllables. The role of intensity is not as consistent. Some (Fry 1955, Beckman 1986) have identified average intensity over the syllable as a possible acoustic correlate of stress, while others (Sluijter & van Heuven 1996b) have argued that spectral tilt (differences in intensity over the frequency spectrum of a given vowel) is more appropriate. Since pitch has a double function that can be represented at the lexical level, as a lexical pitch, or at the sentential level, as an intonation pitch, its double function may also affect its effectiveness as a cue to stress. 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 a question-type intonation. In addition, if stress-contrasted minimal pairs appear in the same position in a sentence in non-focus position, the pitch differences of the minimal pairs are too small to signal the stress (Ladd 1996). It seems that in stress languages, the most common hierarchy of acoustic parameters is duration, intensity and then fundamental frequency in the production of stress while in pitch-accent languages, fundamental frequency is the primary cue to stress.

1.1 Research on Turkic languages

Acoustic analyses of stress have not been extensively undertaken in the Turkic languages (which include the Uyghur language), with the exception of Turkish. Levi (2005) measured duration, intensity and fundamental frequency in near-minimal pairs produced by ten female native Turkish speakers. Levi found that only fundamental frequency was a reliable cue in distinguishing stressed syllables from unstressed syllables in Turkish. As a result, she concluded that Turkish is a pitch-accent language rather than a stress-accent language. While duration and intensity information did also differ across stressed and unstressed syllables, Linear Discriminant Analysis (LDA) revealed that the order of importance was fundamental frequency (90%), intensity (70%) and duration (65%). Using step-wise logistic discriminant analysis, fundamental frequency correctly classified over 99% of the data for verbs and 96% of the data for nouns. Thus, Levi concluded that, in Turkish, only fundamental frequency can be relied upon as a cue for distinguishing stressed syllables from unstressed syllables.

Dobrovolsky (1999) investigated the stress pattern in Chuvash, also a Turkic language. Dobrovolsky did not use minimal pairs; instead he contrasted disyllabic words in four word-stress groups: full–full (in which the first syllable contains a full vowel and the second syllable also contains a full vowel), full–reduced (in which the first syllable contains a full vowel and the second syllable contains a reduced vowel), reduced–full (in which the first syllable contains a reduced vowel and the second syllable contains a full vowel), and reduced–reduced (in which both the first and second syllables contain reduced vowels). Dobrovolsky found that words in Chuvash that have contrasting stress placement are distinguished by duration ratios and overall amplitude and that the duration ratios and overall amplitude are significant cues within each word-stress class except for the reduced–reduced group. In Chuvash, duration and total amplitude (intensity) seem to be critical cues to assigning stress. On the basis of these data, Dobrovolsky claimed that Chuvash is a stress-accent language. Although researchers have categorized Turkish as a pitch-accent language, it seems that some Turkic languages may not be.

For Uyghur, two studies have focused on the acoustic correlates of stress. Liang & Zhang (2008) examined the acoustic correlates of stress in Uyghur using 16 disyllabic words and five newly created non-words produced by ten native Uyghur speakers. The words did not contrast in terms of stress since the authors assumed the stress would fall on the final syllable. Rather, in their analyses, initial syllables were compared to final syllables. The stimuli consisted of four syllable structures (CVCV, CVCCVC, CVCVC and CVCCV). Liang & Zhang (2008) measured the fundamental frequency, duration and peak intensity of both syllables in each word, comparing the first syllables to the second syllables, finding that stressed syllables (always assigned as the final syllable) had longer duration but not higher fundamental frequency or intensity than unstressed syllables (always assigned as the initial syllable). Liang & Zhang (2008) concluded that only duration provided cues to stress due to the fact that in all productions, final syllables were longer than initial syllables, both in real words in Uyghur and their Uyghur non-words. Unfortunately, in this study, word-final lengthening was not controlled for and word stress was also assumed to be in the final position, making the data harder to interpret. In the only other study examining Uyghur stress, Jiang, Liu & Lu (2010) examined disyllabic and trisyllabic words in Uyghur spoken by 12 native Uyghur speakers. They too assumed that stress would fall on the final syllable of the word. They measured duration, total intensity, and average fundamental frequency for each syllable, and they also used ratios of the second syllable to the first syllable. Jiang et al. (2010) found that second syllables were longer in duration (a finding similar to that in Liang & Zhang 2008), and greater in intensity than the first syllables (all ratios were larger than 1). However, they did not provide any statistics. They concluded that both duration and intensity effectively distinguished stressed syllables from unstressed syllables in Uyghur. Given only these two studies and the fact that Uyghur is a Turkic language, with some members of the language group patterning as pitch-accent languages and others as stress languages, a thorough and detailed investigation of Uyghur is needed.

2 Present study

The present study systematically investigated the acoustic correlates of stress in Uyghur, examining duration, intensity and fundamental frequency cues. The research focused on three experiments utilizing minimal pairs as well as disyllabic nouns that contrasted in the first syllables. In addition, Uyghur lexical stress was investigated in a sentential context. The main purpose of this research was to isolate and determine which acoustic parameters provide cues to stress in Uyghur.

Experiment 1 examined minimal pairs in Uyghur which have identical segmental structure but which contrast in terms of stress placement, such as *Acha* /ˈaʧa/ ‘elder sister’ and *aCHA*

Table 1 Minimal pairs (the capitalized syllable has stress) in Uyghur used in Experiment 1.

#	Initial stress	IPA	English gloss	Final stress	IPA	English gloss
1	Acha	/ˈaʧa/	'elder sister'	aCHA	/aˈʧa/	'branching'
2	Ara	/ˈara/	'fork'	aRA	/aˈra/	'between'
3	TÖshük	/ˈtøʃyk/	'kitchen'	töSHÜK	/tøˈʃyk/	'hole'
4	BAla	/ˈbala/	'child'	baLA	/baˈla/	'disaster'
5	CHAtaq	/ˈʧataq/	'bad branch of tree'	chaTAQ	/ʧaˈtaq/	'problem'
6	PAchaq	/ˈpaʧaq/	'leg'	paCHAQ	/paˈʧaq/	'piece'

/aˈʧa/ 'branching'. This experiment allowed a direct examination of stress cues when all segmental information is identical and only stress changes from the first to the second syllable.

Experiment 2 further examined Uyghur stress by contrasting more common disyllabic stimuli, 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, the first syllable is not stressed (the second syllable is stressed), such as *DAka* /ˈdaka/ 'gauze' versus *daLA* /daˈla/ 'plain'.

Experiment 3 examined the acoustic correlates of lexical stress in Uyghur in a sentential context. To control location of the target word in the sentence, declarative assertion sentences and declarative question sentences were contrasted. Word order was identical so that the target word was in exactly the same location, with exactly the same preceding and following context; only the intonation contour (declarative assertion versus declarative question) varied.

These three experimental phonetic studies examining stress correlates in Uyghur will enhance our typological knowledge of stress correlates and will also provide a set of acoustic data for an understudied language.

2.1 Experiment 1

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 are contrasted in terms of stress placement, including six initial stress and six final stress words. Three acoustic parameters were measured: average fundamental frequency, vowel duration, and average intensity. For the analyses, we used ratios of the first syllable to the second syllable (Beckman 1986). If Uyghur is a stress-accent language, fundamental frequency, duration and intensity may be strongly correlated to stress location, similarly to stress cues in English. On the other hand, if Uyghur only uses fundamental frequency as a stress cue, Uyghur will pattern similarly to other pitch-accent languages such as Turkish or Japanese. These data will provide some insight into the stress characteristics of Uyghur as a stress-accent or pitch-accent language.

2.1.1 Method

2.1.1.1 Stimuli

In this experiment, six minimal pairs were used to examine stressed versus unstressed syllables. In Uyghur minimal pairs, stress is contrastive, and thus the meaning is differentiated by the location of stress. For example, consider *Acha* /ˈaʧa/ 'elder sister' and *aCHA* /aˈʧa/ 'branching'. In the first word of the minimal pair, the stress is located on the first syllable, and in the second word of the minimal pair, it is located on the second syllable. All of the stimuli were monomorphemic. The stimuli are shown in Table 1.

All minimal pairs were disyllabic, making it easier to control stress location. In addition, the stimuli were presented to the speakers in appropriate context sentences to provide a

semantic background for each target word. For each token, speakers read a context sentence with the target word, e.g.:[1]

(1) Qeshqer-de öz-i-din chong qiz qérindash-i-ni **acha**
Qeshqer-LOC self-POSS3-ABL big female sibling-POSS3-ACC elder.sister
de-y-du.
say-PRES-3RD.SG
'In Kashger, people called the older female sibling elder sister.'

Speakers then read a fixed sentence context which also included the target word, e.g.:

(2) Men hazir **acha** de-y-men.
I now elder.sister say-PRES-1ST.SG
'Now I say elder sister.'

While the context sentences differed across minimal pairs, the fixed sentence that included the target word was identical across all minimal pairs. The target stimuli in the fixed sentence context were used for all acoustic analyses.

2.1.1.2 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, with two speakers having very little knowledge of English.

2.1.1.3 Procedure

Each subject was recorded individually. Subjects were instructed to read the sentences, presented in the Uyghur Arabic alphabet (*Uyghur Ereb Yéziqi*) that included the context sentence followed by the fixed sentence for each member of a minimal pair. Each participant's speech sounds were recorded with a Marantz PMD 671 solid-state recorder. The speakers were recorded at the University of Kansas, in an anechoic chamber via an Electro-Voice ND767a microphone, placed at a fixed distance from the speaker.

There were three repetitions for each minimal word pair, presented in a random order to the Uyghur speakers. The recordings were digitized at a sampling rate of 22.05 kHz. A total of 180 stimuli were examined (5 speakers × 6 minimal pairs × 2 words × 3 repetitions), 36 tokens for each speaker.

2.1.1.4 Analysis

Three acoustic measurements (fundamental frequency, duration and intensity) were taken for each target word using Praat software (Boersma & Weenink 2011). In all stimuli, vowels were identified from wide-band spectrograms. The beginning of the vowel was the point at which a clear first formant frequency (F1) was observed (onset of F1) and the ending of the vowel was the point where a clear second formant frequency ended (offset of F2). For fundamental frequency, fundamental frequency (in Hz) was measured at 10 ms intervals starting with the onset of the vowel. The first 20% and last 20% of vowel were not used and the average f0 was calculated across the medial interval. For vowel duration, the entire vowel duration (in ms) was measured from onset of F1 to offset of F2. For intensity, the average intensity (in dB) was also obtained based on the entire vowel.

For minimal pairs, the lexically stressed and unstressed syllables were compared using second syllable to first syllable ratios (Beckman 1986).

[1] Abbreviations used in example glosses are as follows: 1ST = first person, 3RD = third person, ABL = ablative, ACC = accusative, ADJZ = an adjectivizer, CL = a clitic, DAT = dative, GEN = genitive, LOC = locative, PL = plural, POSS3 = possessive third person, PRES = present tense, SG = singular.

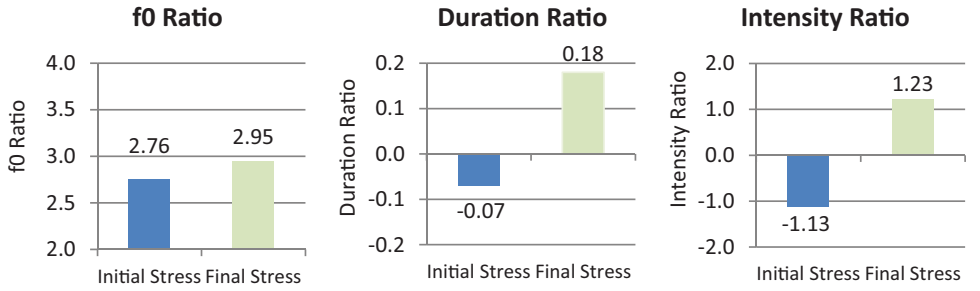

Figure 1 (Colour online) f0 ratios, duration ratios, and intensity ratios for initial and final stress minimal pairs in Experiment 1 (second syllable to first syllable ratios). The dark bar shows the data for stimuli with initial stress and the light bar shows the data for stimuli with final stress.

2.1.2 Results

Repeated Measures ANOVAs were conducted for the ratio values, both by subjects (*F1*) and items (*F2*). Standard repeated measures ANOVAs (utilizing both *F1* and *F2*) are appropriate for balanced designs such as in the present experiments. Separate ANOVAs were conducted for f0, duration and intensity. For f0, the effect of initial versus final stressed words was not significant ($F1(1,4) = 0.21$, $p = .67$; $F2(1,5) = 0.39$, $p = .56$) as shown in Figure 1. The f0 ratio for final stressed words (2.95) was not significantly higher than the f0 ratio for initial stressed words (2.76). f0 ratios did not differentiate the stressed syllables from the unstressed syllables in Uyghur.

For duration, the effect of initial versus final stressed words was significant by items and marginally significant by subjects ($F1(1,4) = 5.34$, $p = .08$; $F2(1,5) = 9.43$, $p = .02$), as shown in Figure 1. The final stressed words (0.18) have a longer duration in the second syllable than the first syllable, and initial stressed words (–0.07) have a shorter duration in the second syllable than the first syllable. Overall, duration does distinguish stress location in Uyghur.

For intensity, the effect of initial versus final stressed words was also significant by items and marginally significant by subjects ($F1(1,4) = 6.22$, $p = .06$; $F2(1,5) = 35.27$, $p = .002$) as shown in Figure 1. The average intensity ratio for final stressed words (1.23) was significantly higher than the average intensity ratio for initial stressed words (–1.13). Intensity is also a cue for distinguishing stressed syllables from unstressed syllables in Uyghur.

2.2 Experiment 2

In Uyghur, minimal pairs contrasting in stress are relatively rare. Some members of the six minimal pairs in Experiment 1 are rarely used and some are only found in regional dialects or in historical usage. For example, *TÖshük* /ˈtøʃyk/, as 'kitchen', is an historical term, and it is currently used only in some geographical 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. Additionally, when both members of the minimal pair originate from Turkic words, some letters or syllables were dropped historically and, as a result, they became minimal pair counterparts. For example, in *Acha* /ˈatʃa/ 'elder sister', the original Turkic version was *aghicha* /aʁiˈtʃa/, which means 'a woman servant in a noble or rich house'. In Experiment 1, five of the six minimal pairs lost some syllables historically and, although they have been part of Uyghur vocabulary for several centuries, they are newer additions to the language. Finally, some members of the minimal pairs were borrowed from Arabic or Farsi. For example, the word *bala* 'disaster' was borrowed from Arabic, between the 9th and the15th centuries, when many Uyghurs converted to Islam and accepted Arabic words via the Koran.

In order to generalize beyond these rare minimal pairs, Experiment 2 examined more common disyllabic words which share the same syllable in 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 pair, the first syllable is not stressed (the second syllable is stressed). For example, in *DAka* 'gauze' versus *daLA* 'plain', the first syllable *da* in *DAka* is stressed, and the first syllable *da* in *daLA* is not stressed. By examining one segmentally-identical syllable (stressed versus unstressed), a greater variety of stimuli could be selected. Experiment 2 thus examines acoustic cues in both stressed and unstressed first syllables in disyllabic words in Uyghur.

One advantage of using disyllabic words in Experiment 2 is to examine the influence of syllable structure in stressed syllables. Little research has been done on the role of syllable weight in Uyghur. Given 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 stress in Uyghur. In order to investigate the role of syllable weight, we examined two distinct disyllabic structures, CVCV and CVCCVC, in Experiment 2.

Saimaiti & Feng (2007) conducted a corpus study examining syllable structure in Uyghur. They used a syllabification algorithm for two corpora: one is the Uyghur Explanatory Dictionary (*Uyghur Tilining Izahliq Lughiti*, Yaqup & Gheyurani 1999), which has 30,169 words (stem types), and the other is a corpus established by Xinjiang University, which has 2,558,810 words. Drawing on these sources, Saimaiti & Feng (2007) categorized the following syllable structures in Uyghur: V, CV, VC, CVC, CVCC, CCV, CCVCC, CVV, and CVVC. The most common syllable structures were CV (38.4%) and CVC (50.33%) in the stem form type and CV (50.01%) and CVC (40.44%) in inflectional type (Saimaiti & Feng 2007). With reference to that study, the most common syllable structures (CV and CVC) were examined in Experiment 2.

In Experiment 2, a new set of stimuli, in which the first syllables contrasted in terms of stress, was used in order to generalize the findings to a more frequently used stimulus set. Additionally, syllable structure (CV versus CVC structure) was investigated to determine whether acoustic cues to stress are affected by syllable weight in Uyghur. If syllable weight has a role in stress, a contrasting pattern across the two types of words (CVCV words and CVCCVC words) may be revealed.

2.2.1 Method

2.2.1.1 Stimuli

Twelve pairs of stimuli were selected, half with CV structure (six pairs of disyllabic CVCV words) and half with CVC structure (six pairs of disyllabic CVCCVC words). All disyllabic stimulus pairs contrasted first syllable stressed words to first syllable unstressed words (i.e. second syllable stressed words), with segmentally identical first syllables.

Given that Uyghur stress is not well understood (explanations invoke vowel length, syllable structure, and location in the stem; see e.g. Yaqup & Gheyurani 1999, Engesaeth, Yakup & Dwyer 2009), a pilot test was conducted to examine perceived stress by native speakers of Uyghur. The purpose of the pilot test was to select the stimuli to be recorded. Since stress is not typically listed in the Uyghur dictionary, the pilot test provided a more objective criterion for stimulus selection. In the test, listeners indicated stress location (first syllable or second syllable stress) and their judgments were the basis for the selection of stimuli for Experiment 2 and Experiment 3.

Seventy-five paired words (Yaqup & Gheyurani 1999) that included 32 CVCV disyllabic word pairs and 43 CVCCVC disyllabic word pairs were initially selected. Word pairs were chosen that had identical segmental content in the first syllable. Additionally, the phonemes /m n l h r ŋ/ were avoided in coda position to ensure accurate segmentation.

Eighteen native Uyghur participants, who had not participated in any of the production experiments, evaluated whether the stress was located on the first or the second syllable for the 150 words from a written list of words. The participants were senior-level psychology students

from Xinjiang Normal University. While the data showed that perceived stress location was not identical across all participants, there was a consistent pattern.

Target stimuli were included when a majority of the Uyghur participants consistently identified stress location for both the first syllable stressed words as well as the second syllable stressed word counterpart (same first syllable segmental content). Six word pairs with CVCV structure (*DAka* /ˈdaka/ 'gauze' – *daLA* /daˈla/ 'plain', *CHAsa* /ˈʧasa/ 'squared' – *chaNA* /ʧaˈna/ 'sledge', *BAza* /ˈbaza/ 'base' – *baHA* /baˈha/ 'price', *DAcha* /ˈdaʧa/ 'villa' – *daDA* /daˈda/ 'father', *DOra* /ˈdora/ 'medicine' – *doQA* /doˈqa/ 'forehead', *POchi* /ˈpoʧi/ 'boaster' – *poTA* /poˈta/ 'western belt') were selected, with each pair including a first syllable stressed word and a first syllable unstressed word. The mean number of participants identifying stress location for the first syllable stressed words (14/18) was similar to the second syllable stressed words (13/18) ($t(5) = 1.965$, $p = .11$, ns). Both members of the target word pairs showed consistent stress location.

The following six word pairs with CVCCVC structure were selected, with each pair including a first syllable stressed word and a first syllable unstressed word: *SEPkün* /ˈsɛpkyn/ 'mole' – *sepDASH* /sɛpˈdaʃ/ 'fellow', *PUTbol* /ˈputbol/ 'football' – *putLASH* /putˈlaʃ/ 'trapping', *KASsir* /ˈkassir/ 'cashier' – *kasTUM* /kasˈtum/ 'suite', *PARnik* /ˈparnik/ 'greenhouse' – *parLASH* /parˈlaʃ/ 'choosing', *BANkir* /ˈbankir/ 'banker' – *banTIK* /banˈtik/ 'tie', and *TOQmaq* /ˈtoqmaq/ 'stick' – *toqQUZ* /toqˈquz/ 'nine'. (In the CVC group, one stimulus, *PUTbol* /ˈputbol/ 'football' and its counterpart *putLASH* /putˈlaʃ/ 'trapping' were eliminated from the data analysis. The devoicing of [u] occurred for all subjects and the vowel could not be isolated. Therefore, only five word pairs in the CVC group were analyzed.) The mean number of participants identifying stress location for the first syllable stressed words (12/18) was not significantly different from the second syllable stressed words (11/18) ($t(5) = 1.12$, $p = .31$, ns). Both members of the target word pairs received consistent stress location.

Comparing the CV to the CVC stimuli, Uyghur participants in general were more consistent in assigning stress to the CV structure stimuli (14/18) compared to the CVC structure stimuli (12/18) ($t(5) = 4.44$, $p = .007$).

For Experiment 2, then, first syllable stressed disyllabic words were compared to first syllable unstressed words, examining segmentally identical first syllables. In addition, different syllabic structures (CV and CVC) were contrasted.

2.2.1.2 Participants

Seven male native Uyghur speakers (aged 24–35 years old) living in the USA were recorded individually. None had participated in Experiment 1 and none of them had speech or hearing disorders. All were also fluent in Mandarin Chinese and all of them had some knowledge of English. They learned English as a third language at the age 16 (age of arrival (AOA) in the USA: 13–18 years). All of them were graduate students at universities in the USA.

2.2.1.3 Procedure

Each subject was recorded individually. Each participant's utterances were recorded with a Marantz PMD 671 solid-state recorder via an Electro-Voice ND767a microphone, placed at a fixed distance from the speaker, in a quiet room. As in Experiment 1, subjects were instructed to read the sentences that included a context sentence (to provide a semantic background for each target word) followed by the fixed sentence for each stimulus pair. The target stimuli in the fixed sentence context were used for all acoustic analyses.

There were three repetitions for each word pair, presented in a random order to the Uyghur speakers. The recordings were digitized at a sampling rate of 22.05 kHz. A total of 504 stimuli were examined (7 speakers × 12 pairs × 2 words × 3 repetitions), 72 tokens for each speaker.

2.2.1.4 Analysis

As in Experiment 1, acoustic measurements were collected. Mean vowel fundamental frequency, vowel duration and mean vowel intensity measures of the first syllable were examined. The initial syllable of first syllable stressed words was directly compared to the

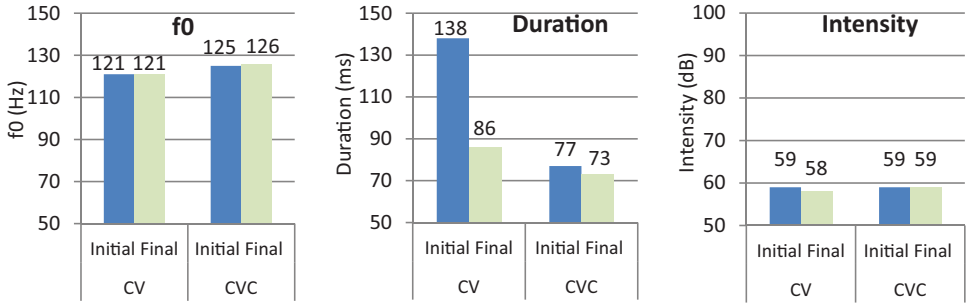

Figure 2 (Colour online) Average fundamental frequency (Hz), average duration (ms), and average intensity (dB) across two types of syllabic structure (CV versus CVC) are shown. The dark bar shows the data for stimuli with initial stress and the light bar shows the data for stimuli with final stress.

initial syllable of first syllable unstressed stimuli (for example, comparing the first syllable in *DAka* 'gauze' with the first syllable of *daLA* 'plain').

2.2.3 Results

A two-way Repeated Measure ANOVA (Stress × Syllable Structure) was conducted for fundamental frequency, duration and intensity measures.

For f0, the main effect of Stress was not significant ($F1(1,6) = 0.44$, $p = .53$; $F2(1,9) = 0.44$, $p = .52$). As shown in Figure 2, no significant differences were found across stressed and unstressed syllables. For f0, the main effect of Syllable Structure was significant ($F1(1,6) = 19.13$, $p = .005$; $F2(1,9) = 5.72$, $p = .04$). The CVC stimuli (126 Hz) had higher f0 on average than the CV stimuli (121 Hz). The interaction was not significant across both subject and item analyses ($F1(1,6) = 15.15$, $p = .008$; $F2(1,9) = 1.76$, $p = .218$).

For duration, the main effect of Stress was significant ($F1(1,6) = 50.41$, $p = .0001$; $F2(1,9) = 34.68$, $p = .0001$), with the duration of stressed syllables (108 ms) significantly longer on average than unstressed syllables (80 ms). As shown in Figure 2, duration provides information about stress location in disyllabic nouns.

As expected, the main effect of Syllable Structure was also significant ($F1(1,6) = 82.85$, $p = .0001$; $F2(1,9) = 26.62$, $p = .001$). Overall, vowels in the open CV syllables (112 ms) were longer on average than vowels in the closed CVC syllables (75 ms). The interaction was also significant ($F1(1,6) = 56.02$, $p = .0001$; $F2(1,9) = 27.18$, $p = .001$). Paired *t*-tests for the CV stimuli revealed a significant difference between stressed syllables (138 ms) and unstressed syllables (86 ms) ($p = .0001$) while for the CVC stimuli, only a marginally significant difference ($p = .076$) was found between the stressed syllables (77 ms) and the unstressed syllables (73 ms).

Overall, duration differences between stressed and unstressed syllables were greater for the CV syllable stimuli than for the CVC syllable stimuli, with the CV stimuli showing more systematic differences between stressed versus unstressed syllables than the CVC stimuli.

For intensity, the main effect of Stress was not significant ($F1(1,6) = 0.99$, $p = .357$; $F2(1,9) = 1.40$, $p = .267$), as shown in Figure 2. There was no significant difference between the stressed syllables (59 dB) and the unstressed syllables (58 dB) in terms of intensity. The main effect of Syllable Structure was also not significant ($F1(1,6) = 0.12$, $p = .745$; $F2(1,9) = 0.14$, $p = .72$). The interaction was not significant across both subject and item analyses ($F1(1,6) = 5.64$, $p = .055$; $F2(1,9) = 3.00$, $p = .117$).

Overall, for these disyllabic stimuli, the data show that intensity does not differentiate the stressed syllables from the unstressed syllables.

2.3 Experiment 3

When speakers perceive the prominence of words or syllables in sentences, sentential intonation may overlay and interact with perception of lexical stress. While lexical stress can be observed by changes in fundamental frequency, duration or intensity, sentence intonation variations can also be indicated by similar acoustic cues (Beckman 1986). Given this, cues to lexical stress may change depending not only on the position of the target word in the sentence, but also on the overall intonation contour. In many languages, a raised pitch at the end of a sentence indicates that an utterance is intended as a question compared to a lowered pitch that indicates a declarative sentence (Lieberman 1967, Hirst & Di Cristo 1998).

Experiment 3 investigated the interaction of sentential intonation with lexical stress. Declarative and question intonation were compared. To control location of the target word in the sentence, declarative assertion sentences and declarative question sentences were contrasted. In such sentences, word order was identical so that the target word was in exactly the same location and with exactly the same preceding and following context; only the intonation contour (declarative assertion and declarative question) varied. These two types of utterance have clearly distinct intonation contours in Uyghur, with declarative assertion sentences having falling intonation patterns and declarative question sentences having rising final intonation patterns. While the exact details of these contours are a matter of further research, the different pitch contours may interact with the acoustic correlates of stress of words that are in sentence-final position in the two kinds of sentences.

Lindström & Remijsen (2005) examined the Kuot language in which they embedded stimuli in four distinct utterance positions to determine whether f0 cues may play a role in different sentential positions. They found that the cues to lexical stress remained the same regardless of sentence position. While duration and vowel quality are cues to stress in Kuot, no interaction was found between intonation contour and cues to stress. For Kuot, f0 at the sentential intonation level was separated from f0 at the lexical level. However, Hirst & Di Cristo (1998) had suggested that f0 might play a greater role (especially at the end of the sentence when f0 may be rising most prominently) and possibly combine with duration and intensity cues to differentiate stressed syllables from unstressed syllables at the lexical level, especially in final position.

In Experiment 3, the interaction of lexical stress with sentential intonation using declarative sentences as assertions and declarative sentences as questions was examined in Uyghur. In Experiment 1 and Experiment 2, f0 did not provide a cue to stress in Uyghur. In Experiment 1, duration and intensity were stress cues, and in Experiment 2, duration was a stress cue. Experiment 3 provides a window for testing the interaction of the lexical and sentential level intonation cues in Uyghur, by examining lexically contrastive words in the context of distinct f0 intonational contours (declarative and question sentential contexts). Two outcomes are possible: (i) an enhancement effect that indicates the acoustic parameters (f0, duration and intensity) are increased by changes in sentential intonation; and (ii) an attenuation effect that indicates that the acoustic parameters (f0, duration and intensity) are minimized by the changes in sentential intonation.

2.3.1 Method

2.3.1.1 Stimuli

The six word pair CV stimuli from Experiment 2 were used. Two sentences were selected: one was a declarative sentence with assertion (DA) and another was a declarative sentence with question (DQ). In English, the sentence *Katrina bought a book* is a declarative sentence with assertion; on the other hand, the sentence *Katrina bought a book?* is a declarative sentence with question. In Uyghur, declarative assertion and declarative question sentences are similar to English. Two types of sentences were used, contrasting disyllabic nouns in sentence-final position. Therefore, for the six word pair stimuli (with first syllable and second syllable stress), there were 72 sentence tokens (6 pairs × 2 words × 2 sentence conditions × 3 repetitions)

for each participant. The total was 432 tokens (6 speakers × 6 pairs × 2 words × 2 sentence conditions × 3 repetitions).

2.3.1.2 Participants

Six male native Uyghur speakers who participated in Experiment 2 were recorded individually.

2.3.1.3 Procedure

Each word was recorded in two sentential conditions: declarative assertion (DA; see example sentences (3) and (4) below) and declarative question (DQ; see example sentences (5) and (6)). Identical sentences were used in both conditions.

(3) Biz-ge kérek bol-ghan-i chong we pakiz
we-DAT necessary become-ADJZ-POSS3 big and clean
daka. (declarative assertion)
gauze
'What we need is big and clean gauze.'

(4) U-ning yil-lar-che siz-ghan resim-lar-i güzel bir
s/he-GEN year-PL-CL draw-ADJZ painting-PL-POSS3 beautiful a
dala. (declarative assertion)
countryside
'The paintings she draws every year are of a beautiful countryside.'

(5) Biz-ge kérek bol-ghan-i chong we pakiz
we-DAT necessary become-ADJZ-POSS3 big and clean
daka? (declarative question)
gauze
'What we need is big and clean gauze?'

(6) U-ning yil-lar-che siz-ghan resim-lar-i güzel bir
s/he-GEN year-PL-CL draw-ADJZ painting-PL-POSS3 beautiful a
dala? (declarative question)
countryside
'The paintings she draws every year are of a beautiful countryside?'

All sentences were controlled for the number of syllables. For Experiment 3, the target word was in final position for both the declarative assertion and declarative question sentence types. Sentences were randomized for recording and three repetitions were elicited.

2.3.1.4 Analysis

Measurements used in Experiment 3 were similar to those in Experiment 2.

2.3.2 Results

A two-way Repeated Measures ANOVA (Stress × Sentence Type) was conducted for fundamental frequency, duration and intensity measures, in which Stress (first syllable stressed versus first syllable unstressed) and Sentence Type (DA versus DQ) were within-subject variables for both the subject and item analysis.

For f0, no significant effect of Stress was found ($F1(1,5) = 1.01, p = .361$; $F2(1,5) = 5.29, p = .07$), as shown in Figure 3. The significant main effect of Sentence Type ($F1(1,5) = 4.20, p = .096$; $F2(1,5) = 47.55, p = .001$) indicated that DQ sentences (125 Hz) had an overall higher f0 on average than the DA sentences (118 Hz). The interaction was not significant ($F1(1,5) = 4.12, p = .098$; $F2(1,5) = 1.49, p = .276$). This non-significant interaction indicated that there were no differences in use of f0 across sentential type. The overall f0 main effects were due to sentential intonation rather than lexical differences.

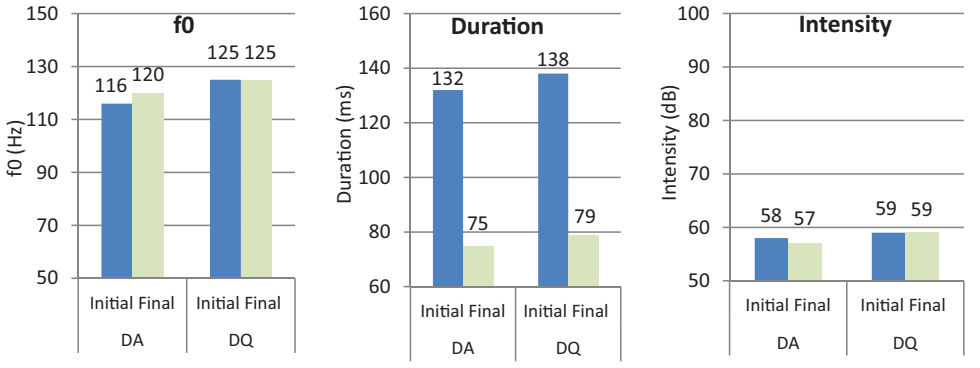

Figure 3 (Colour online) Average fundamental frequency (Hz), average duration (ms), and average intensity (dB) across the two types of sentence contexts: declarative assertion (DA) and declarative question (DQ). The dark bar shows the data for stimuli with initial stress and the light bar shows the data for stimuli with final stress.

Overall, these results indicate that speakers raised the f0 sentence-finally, regardless of stress location, due to the declarative question intonation. There was no enhancement of f0 cues in contrastive sentential contexts. f0 is not used as a cue to stress location, even in highly-marked contexts such as a question intonation.

For duration, a significant main effect of Stress ($F1(1,5) = 101.74$, $p = .0001$; $F2(1, 5) = 35.81$, $p = .002$) indicated that the stressed syllables (135 ms) had a significantly longer duration on average than the unstressed syllables (77 ms), as shown in Figure 3. Duration is consistently used as a cue to stress location in Uyghur. There was no significant main effect for Sentence Type ($F1(1,5) = 1.39$, $p = .292$; $F2(1,5) = 37.69$, $p = .002$), suggesting that vowel duration in the declarative question sentences (109 ms) was not different from vowel duration in the declarative assertion sentences (104 ms). In addition, the interaction was not significant ($F1(1,5) = 0.70$, $p = .44$; $F2(1,5) = 0.87$, $p = .393$). The obtained duration effect was not due to sentential intonation but due to lexical stress.

Overall, for duration, there was neither enhancement nor attenuation of duration cues in contrastive sentential contexts. In other words, the changes in duration from the stressed to the unstressed syllables in DA sentences were similar to the changes observed in DQ sentences.

For intensity, no significant main effect of Stress ($F1(1,5) = 3.11$, $p = .138$; $F2(1,5) = 2.47$, $p = .177$) was found, indicating that the intensity in the stressed syllables (59 dB) was not significantly greater than that in the unstressed syllables (58 dB), as shown in Figure 3. There was also no significant effect of Sentence Type ($F1(1,5) = 0.57$, $p = .484$; $F2(1,5) = 4.84$, $p = .079$), signifying that overall intensity in declarative question sentences (59 dB) was not greater than in declarative assertion sentences (58 dB). The interaction was also not significant ($F1(1,5) = 0.002$, $p = .962$; $F2(1,5) = 0.034$, $p = .862$).

Overall, for intensity, there was no enhancement or attenuation of intensity cue in contrastive sentential contexts. The results indicated that the participants' behavior in both the stressed syllables and the unstressed syllables was similar for DA and DQ sentential intonations.

3. General discussion and conclusion

The present research provides an acoustic analysis of lexical stress patterns in Uyghur, an understudied Turkic language. The basic research question was to identify which acoustic parameters were correlated with stress location in Uyghur. Three experiments were conducted

investigating the acoustic correlates of stress in Uyghur. In all experiments, three distinct acoustic correlates were analyzed: fundamental frequency, duration and intensity.

Experiment 1 focused on minimal pairs in a fixed sentence condition. Using a small set of minimal pairs, we found that only duration and intensity of the vowel were consistent cues to distinguishing stressed syllables from unstressed syllables. f0 was not a stress cue in Uyghur.

Experiment 2 expanded this investigation, examining a wider variety of disyllabic nouns that contrasted in the first syllable in terms of stress. In addition, syllable structure (CV versus CVC) was manipulated to determine whether acoustic cues to stress are affected by syllable weight in Uyghur. As in Experiment 1, fundamental frequency, duration and intensity of the vowels were measured. For this more representative and expanded set of stimuli, Experiment 2 clearly showed that vowel duration was a strong cue to stress location while f0 and intensity did not distinguish stressed from unstressed syllables. The present data also reveal an interesting aspect about the perception of Uyghur stress. When we asked participants to indicate stress location (perception experiment for stimulus selection in Experiment 2), stress judgments were not entirely reliable. While a majority of listeners did agree on stress location, these decisions were not identical across all Uyghur listeners. While, perceptually, listeners are ambivalent about where stress is located, in production they consistently indicate stress using duration cues.

Finally, in Experiment 2, participants were sensitive to syllable structure in their productions of the Uyghur stimuli. While vowels in open syllables (CV) were longer than in closed syllables (CVC), an expected finding, we also found that duration differences between stressed and unstressed syllables were greater for the CV syllable stimuli than for the CVC syllable stimuli. Interestingly, although CVC syllable structure was more complex than CV syllable structure, CV syllables showed stronger duration effects in terms of stress as compared to their CVC counterparts.

A comparison of the present data with those in the two other studies examining Uyghur reveals a similar pattern. Liang & Zhang (2008) examined the stress pattern in Uyghur using both disyllabic words and newly created non-words while Jiang et al. (2010) investigated the stress pattern in Uyghur, in disyllabic and trisyllabic words with varying syllable structure. Drawing on data similar to the present data, Liang & Zhang (2008) concluded that duration could distinguish the stressed syllables from the unstressed syllables and that fundamental frequency and intensity were inconsistent in indicating stress. Jiang et al. (2010) additionally found intensity differences but provided only descriptive analyses to support the claims. However, there were a number of methodological concerns with these studies. The observed duration effects of Liang & Zhang (2008) and the absence of fundamental frequency cues could have been due to phrase-final lengthening and a concomitant f0 rise in the final position since they compared final syllables to initial syllables. Unlike their study, we used minimal pairs or disyllables that minimally contrasted in their first syllable in order to avoid the confounding effect of final lengthening or f0 rise in the final position. Despite these issues, similar results were obtained across both these studies in terms of the role of duration and fundamental frequency. However, differences were found with the current study in that while Liang & Zhang (2008) as well as Jiang et al. (2010) observed that the duration effects were the same across different syllable structures and concluded that syllable structure did not influence stress pattern, neither study contrasted structure in controlled comparisons and measures often included all combinations of V, CV, and CVC syllable structures. By systematically examining minimal pairs and comparing syllables which differ only in stress location, the present data show subtle but systematic differences in syllable structure that influence acoustic cues to stress in Uyghur.

Experiment 3 focused on the interaction between lexical stress and sentential intonation, investigating whether f0 plays a role in lexical stress when f0 is also used to indicate intonation. In Experiment 3, declarative assertion and declarative question sentences were used, because we could control the position of the target word in the final position of both sentence types. Examining the CV stimuli, we again found that duration, but not intensity or f0, was a cue to

stress location. The lack of a significant interaction between stress and sentential intonation indicated that the obtained duration effects resulted from lexical stress rather than from the contrasting sentential intonation.

Interestingly, the present data show that the intonation contour has little effect on lexical stress cues. The contrasting intonational context did not enhance or attenuate the differences between stressed and unstressed syllables nor did it affect the specific cues used to signal stress in Uyghur, a possible reason why Liang & Zhang (2008) and Jiang et al. (2010) were so consistent in their observation of duration as one of the most robust cues to stress in Uyghur. Previous studies have shown that observed stress cues, especially fundamental frequency and intensity, are influenced by accent location (Sluijter & van Heuven 1996a, b). In the present study, accent location was not systematically manipulated. However, the three experiments presented here do examine lexical stress in a number of different sentence locations, in both non-final and final positions, as well as in contrasting sentential contexts. Nevertheless, further research systematically examining sentential accent is needed to precisely clarify the relationship between stress and accent in Uyghur.

Overall, then, the present three experiments consistently show that duration is a robust acoustic cue to stress in Uyghur while both f0 and intensity do not play a substantial role in distinguishing stressed syllables from unstressed syllables in accented position. Uyghur does not use pitch cues when distinguishing stressed syllables from unstressed syllables.

The present data directly address the debate that has focused on whether Uyghur, a Turkic language, is a pitch-accent language similar to Turkish, or whether Uyghur is a stress language similar to English. While the present data are based on a small set of participants, the three experiments examining the acoustic correlates of stress in Uyghur native speakers consistently show that duration cues, and not fundamental frequency or intensity, indicate stressed syllables in Uyghur. Duration clearly distinguishes stressed syllables from unstressed syllables in Uyghur. As such, Uyghur patterns very similarly to stress languages and patterns very differently from pitch-accent languages.

Unlike Levi's (2005) findings that Turkish used f0 as the only cue to assigning stress location, Uyghur used duration for assigning stress location, not f0 or 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 in the stimuli used in the research. In the Turkish study, the words were at least trisyllabic words rather than 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) examined disyllabic nouns in Chuvash, another Turkic language. 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 (the first syllable has a full vowel and the second syllable also has a full vowel), full–reduced, reduced–full, and reduced–reduced disyllables. By measuring duration and intensity, Dobrovolsky found that the duration ratio and total amplitude were significant cues within each word-stress class except for the reduced–reduced group, which was considered as having no stress. Even though Dobrovolsky used different stress groups, the result confirmed that duration and intensity were cues to stress for Chuvash. Interestingly, while in Chuvash intensity was also of some significance, the role of intensity was not as strong as that of duration according to Dobrovolsky.

In Turkic language studies on lexical stress, duration is a robust cue to stress in Uyghur and Chuvash, and it could be a cue for Turkish too if we focus on the results with absolute values. The role of intensity was mixed across studies, although its contribution is less than that of duration. While f0 was not tested in Chuvash, it was found to be a cue in two studies (of Turkish by Levi 2005 and of Uyghur by Liang & Zhang 2008). However, our results as

well as the results from Jiang et al. (2010) in Uyghur do not show an f0 effect. The present results clearly and systematically show that f0 is not a cue to stress in Uyghur. Uyghur does not use f0 to distinguish stressed syllables from unstressed syllables. Therefore, we propose that Uyghur does not pattern like a pitch-accent language, but rather like a stress language that uses duration to cue stress.

Acknowledgements

The authors would like to thank Allard Jongman, Jie Zhang, Arienne Dwyer, and three anonymous reviewers for their insightful and constructive comments on earlier versions of the manuscript. The authors would also like to thank Meghripet Kenji for help in data collection. Part of this research is based on a doctoral dissertation by the first author at the University of Kansas.

References

Altmann, Heidi. 2006. *Perception and production of second-languages stress: A cross-linguistic experimental study*. Ph.D. dissertation, University of Delaware, Newark, Delaware.

Beckman, Mary E. 1986. *Stress and non-stress accent*. Berlin: de Gruyter Mouton.

Beckman, Mary E. & Jan Edwards. 1994. Articulatory evidence for differentiating stress categories. In Patricia A. Keating (ed.), *Papers in Laboratory Phonology III: Phonological structure and phonetic form*, 7–33. Cambridge: Cambridge University Press.

Boersma, Paul & David Weenink. 2011. Praat: Doing phonetics by computer (version 5.3.03). http://www.praat.org/ (accessed 21 November 2011).

Chen, Yang, Michael Robb, Harvey Gilbert & Jay Lerman. 2001. Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics* 15(6), 427–440.

Dobrovolsky, Michael. 1999. The phonetics of Chuvash stress: Implications for phonology. *14th International Congress of Phonetic Sciences* (ICPhS XIV), Berkeley, CA, vol. 1, 539–542.

Engesæth, Tarjei, Mahire Yakup & Arienne Dwyer. 2009. *Greetings from Teklimakan: A handbook of modern Uyghur*, vol. 1. Lawrence, KS: KU [Kansas University] Scholar Works.

Flege, James E. & James Hillenbrand. 1984. Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America* 76(3), 708–721.

Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27(4),765–768.

Fry, Dennis B. 1958. Experiments in the perception of stress. *Language and Speech* 1(2), 126–152.

Gay, Thomas. 1978. Physiological and acoustic correlates of perceived stress. *Language and Speech* 21(4), 347–353.

Gordon, Matthew. 2004. A phonological and phonetic study of word-level stress in Chickasaw. *International Journal of American Linguistics* 70(1), 1–32.

Hirst, Daniel & Albert Di Cristo. 1998. A survey of intonation systems. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 1–44. Cambridge: Cambridge University Press.

Huss, Volker. 1978. English word stress in the post-nuclear position. *Phonetica* 35, 86–105.

Hyman, Larry M. 2009. How (not) to do phonological typology: The case of pitch accent. *Language Sciences* 31, 213–238.

Jiang, Haiyan, Yan Liu & Li Lu. 2010. 维吾尔语词重音实验研究《民族语文》[An experimental study of stress in Uyghur words]. 第三期 [Ethnic study] 3, 67–71.

Juffs, Alan. 1990. Tone, syllable structure and inter language phonology: Chinese learners' stress errors. *International Journal of Applied Linguistics* 2(1), 99–117.

Kochanski, Greg, Esther Grabe, John Coleman & Burton Rosner. 2005. Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America* 118(2), 1038–1054.

Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Lai, Yuwen. 2008. *Acoustical realization and perception of English lexical stress by Mandarin learners*. Ph.D. dissertation, University of Kansas.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.

Levi, Susannah V. 2005. Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association* 35, 73–97.

Liang, Jie & Qing Zhang. 2008. 维吾尔语双音节词重音的产出实验研究 [A production study on stress pattern in disyllabic words in the Uyghur language].《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》[Proceedings of the 8th Phonetics Conference of China (PCC 2008) and the International Symposium on Phonetic Frontiers (ISPF)].

Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America* 33, 451–454.

Lieberman, Philip. 1967. *Intonation, perception, and language*. Cambridge, MA: MIT Press.

Lindström, Eva & Bert Remijsen. 2005. Aspects of prosody of Kuot, a language where intonation ignores stress. *Linguistics* 43(4), 839–870.

Morton, John & Wiktor Jassem. 1965. Acoustic correlates of stress. *Language and Speech* 8(3), 159–181.

Nakatani, Lloyd H. & Carletta H. Aston. 1978. Perceiving stress patterns of words in sentences. *Journal of the Acoustical Society of America* 63, S55.

Nash, Carlos M. 2005. A phonetic and phonological approach to stress in Sinhala verbs. *Workshop on Sinhala Linguistics*, University of California, Santa Barbara, 4–5 June. http://www.linguistics.ucsb.edu/research/cmnash_2005_stress_verb.pdf.

Ortega-Llebaria, Marta & Pilar Prieto. 2011. Acoustic correlates of stress in Central Catalan and Castilian Spanish. *Language and Speech* 54(1), 73–97.

Recasens, Daniel & Aina Espinosa. 2006. Dispersion and variability of Catalan vowels. *Speech Communication* 48, 645–666.

Saimaiti, Maimati & Zhiwei Feng. 2007. A syllabification algorithm and syllable statistics of written Uyghur. *4th Corpus Linguistics Conference*, Birmingham, UK. http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2007/153Paper.pdf.

Sereno, Joan A. 1986. Stress pattern differentiation of form class in English. *Journal of the Acoustical Society of America* 79, S36.

Sereno, Joan A. & Allard Jongman. 1995. Acoustic correlates of grammatical class. *Language and Speech* 38, 57–76.

Sluijter, Agaath M. C. & Vincent J. van Heuven. 1996a. Acoustic correlates of linguistic stress and accent in Dutch and American English. *4th International Conference on Spoken Language Processing*, Philadelphia, PA, 630–633.

Sluijter, Agaath M. C. & Vincent J. van Heuven. 1996b. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100(4), 2471–2485.

Turk, Alice E. & James R. Sawusch. 1996. The processing of duration and intensity cues to prominence. *Journal of the Acoustical Society of America* 99(6), 3782–3790.

Yaqup, Abliz & Ghenizat Gheyurani. 1999. *Uyghur Tilining Izahliq Lughiti* [Uyghur explanatory dictionary]. Urumqi: Xinjang Xelg Neshriyati [Xinjiang People's Publishing House].

Zhang, Yanhong & Alexander L. Francis. 2010. The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics* 38(2), 260–271.

Zhang, Yanhong, Shawn L. Nissen & Alexander L. Francis. 2008. Acoustic characteristics of English lexical stress produced by Mandarin speakers. *Journal of the Acoustical Society of America* 123(6), 4498–4513.