

## A phonetic study of length and duration in Kyrgyz vowels

Nathaniel Ziv Stern & Jonathan North Washington\*

**Abstract.** This paper examines the phonetic correlates of the (phonological) vowel length contrast in Kyrgyz to address a range of questions about the nature of this contrast, and also explores factors that affect (phonetic) duration in short vowels. Measurement and analysis of the vowels confirms that there is indeed a significant duration distinction between the Kyrgyz vowel categories referred to as short and long vowels. Preliminary mid-point formant measurements show that there may be some accompanying spectral component to the length contrast for certain vowels, but findings are not conclusive. A comparison of F0 dynamics and spectral dynamics through long and short vowels does not yield evidence that some long vowels may in fact be two heterosyllabic short vowels. Analysis shows that duration is associated with a vowel’s presence in word-edge syllables in Kyrgyz, as anticipated based on descriptions of word-final stress and initial prominence. However, high vowels and non-high vowels are found to consistently exhibit opposite durational effects. Specifically, high vowels in word-edge syllables are longer than high vowels in medial syllables, while non-high vowels in word-edge syllables are shorter than non-high vowels in medial syllables. This suggests either a phenomenon of durational neutralisation at word edges or the exaggeration of durational differences word-medially, and is not taken as a case of word-edge strengthening. Proposals for how to select from between these hypotheses in future work are discussed.

**Keywords.** Turkic languages; acoustic phonetics; Kyrgyz; vowel duration; vowel length; edge effects

**1. Introduction.** This paper sets out to answer two basic questions about Kyrgyz: (1) is the vowel length distinction purely one of duration, and (2) what else affects the duration of vowels in Kyrgyz? In terms of the vowel length distinction, we examine (a) whether short and long vowels have distinct durations, (b) whether the spectral properties of long and short vowel pairs may also be different, and (c) whether it’s possible that long vowels may in fact be two heterosyllabic (“dual-headed”) short vowels [V.V] instead of singleton long vowels [V:]. To do this, we measured and analysed, respectively, (a) vowel duration, (b) the values of the first and second formants at vowel midpoints, and (c) the spectral dynamics and pitch (F0) dynamics of vowels. Results are presented in (4.3), (4.1), and (4.2), respectively. We found (a) that the durational difference between short and long vowels is significant, confirming previous findings; (b) that there are possible spectral differences between short and long vowels; and (c) that there is not good evidence to suggest an analysis of long vowels as dual-headed.

In terms of other confounds of duration, we examined various factors which are generally known to have the potential to correlate with vowel duration, such as syllable number, syllable position, stress, voicing of following consonant, and whether a tautosyllabic consonant follows. Tests were performed on short vowels only due to limitations of the data set regarding long vowels. We confirmed previous findings that high vowels are significantly shorter than non-high vowels (4.3). We found no consistent

---

\* This research was made possible by an Indiana University Grant-in-Aid of Doctoral Research, Swarthmore Faculty Support Grants, and a Joel Dean Summer Research Fellowship. We would like to recognise the participation of two anonymous Kyrgyz-speaking consultants and thank them for their time. We also very much appreciate the comments of several anonymous reviewers and the feedback we received from the participants of Tu+4, especially Michael Daniel and Iliia Chechuro. Special thanks also to Steve Wang for consultation on matters of statistics, Dylan Charter and Harper Pollio-Barbee for their time assisting with annotation, and Tolgonay Kubatova for periodic consultation about Kyrgyz. All errors in data, methods, and interpretation are our own. Authors: Nathaniel Ziv Stern ([nzivstern@gmail.com](mailto:nzivstern@gmail.com)) & Jonathan Washington ([jonathan.washington@swarthmore.edu](mailto:jonathan.washington@swarthmore.edu)), Swarthmore College.

effect of whether the following consonant was voiced or of whether the vowel was in an open syllable. We did find a strong effect of syllable position—that is, whether the syllable a vowel is in is an initial, medial (non-initial and non-final), or final syllable of a word—for nearly all vowels (4.4). For the most part, vowels in word-edge syllables—that is, initial and final syllables—patterned together. Somewhat surprisingly, however, opposite effects were seen for high and non-high vowels, with the former being longer in word-edge syllables and the latter being shorter in word-edge syllables (4.5).

We propose two hypotheses for how to understand this last finding, both seemingly at odds with the literature on strengthening at prosodic edges, and propose ways that they may be chosen between in future work.

Section 2 provides background on these questions, section 3 outlines our methodology, section 4 presents and interprets the results, and section 5 concludes and offers paths for the continuing pursuit of these questions.

**2. Background.** Evidence for contrastive vowel length in Kyrgyz is presented in section 2.1. Section 2.3 explores the different possible ways long vowels could be structured in Kyrgyz, while section 2.4 overviews other phonetic phenomena that could be active in the vowel length contrast and associated with duration.

**2.1. CONTRASTIVE VOWEL LENGTH IN KYRGYZ.** Kyrgyz can be described as exhibiting contrastive vowel length, demonstrated by minimal pairs like those in (1-3).

- |     |  |     |  |     |  |
|-----|--|-----|--|-----|--|
| (1) | a. [uɫu] ‘howl’<br>b. [uɫu:] ‘older’<br>c. [u:ɫu] ‘son–POSS.3’<br>d. [u:ɫu:] ‘poisonous’ | (2) | a. [q <sup>h</sup> uru] ‘belt–POSS.3’<br>b. [q <sup>h</sup> uru:] ‘build–VN’<br>c. [q <sup>h</sup> u.ru] ‘fry’<br>d. [q <sup>h</sup> u.ru:] ‘fry–VN’ | (3) | a. [t <sup>h</sup> oq] ‘full (sated)’<br>b. [t <sup>h</sup> o:q] ‘chicken’ |
|-----|--|-----|--|-----|--|

Additional minimal pairs that may be used to establish the length distinction for short/long vowel pairs not demonstrated above are [syty] ‘milk–ACC’ and [syty:] ‘milky’, [er] ‘husband’ and [e:r] ‘saddle’, [søk] ‘swear/curse’ and [sø:k] ‘bone’, and [sat] ‘sell’ and [sa:t] ‘hour, clock’. These pairs establish that vowel length is not dependent on position and that it bears some functional load. The entire set of short vowel phonemes of standard (northern) Kyrgyz is /i/, /y/, /e/, /ø/, /a/, /o/, /u/, and /u:/, and the generally recognised long vowel phonemes are /y:/, /e:/, /ø:/, /a:/, /o:/, and /u:/. The two “missing” long correspondences (to the short vowels /i/ and /u/) are discussed in section 2.2.

**2.2. NOTE ON LENGTH AND HIGH UNROUNDED VOWELS.** Not all Kyrgyz short vowels have a corresponding long vowel; specifically, whether there exist long counterparts of /i/ and /u/ in Kyrgyz is a much debated topic. As overviewed by Imart (1981: §196–§212), sources disagree as to whether the sequences /ij/ and /iji/ and /uj/ and /ujju/ are phonetically realised as long versions of /i/ and /u/ and as to the phonological status of these sequences as /i:/ and /u:/.

In our consultant’s speech, we found that some utterances of the above sequences may be realised with a glide and that some are realised more like [i:] and [i:], respectively. It is possible that this indicates phonologisation of long vowels from these sequences is in progress (perhaps forming long counterparts to /i/ and /u/, or perhaps not); e.g., pronunciations with glides may be more careful pronunciations, reflecting the phonological form, or “read speech” pronunciations, reflecting influence from the orthographic form.

**2.3. LONG VOWELS VERSUS DUAL-HEADED VOWELS.** There are open questions as to how the phonological length distinction in Kyrgyz is implemented phonetically.

Some Turkic languages have been characterised as having “dual-headed” or “non-contiguous” long vowels (e.g., Shor, cf. Упретешев 2017), which differ from normal long vowels [V:] in that they function

as heterosyllabic [V.V] sequences in hiatus. In the case of Kyrgyz, at least one source has suggested the presence of dual-headed vowels: Батманов (1938: 17) states that, in his opinion, “long *aa* ... generally does not exist in the Kyrgyz language: in the words *şaar*, *maal*, etc., we have two syllables (=two independent, but identical vowels)” (translation ours; transcription left unmodified; glosses of examples: ‘city’, ‘time’). Батманов (1939: 39) further states that he perceives two distinct [a] sounds in the middle of roots with /a:/, like /sa:t/ ‘hour, clock’. Nothing more appears to be made of these claim, and other long vowels are treated as single-headed vowels in the rest of Батманов’s work.

It has also been suggested that some languages may contrast these two types of vowels (Ladefoged & Maddieson 1996: 320–321). For example, it has been reported that Kikamba exhibits a contrast between short vowels that have coalesced to a single long vowel [V:] and two short vowels of the same quality [V.V]; the two “types of long vowel” may be distinguished both by phonological characteristics and by virtue of the fact that a single long vowel is reported to have a much shorter duration than two short vowels—an average of 127 milliseconds as opposed to 232 milliseconds, respectively (Roberts-Kohn 1995: 317–318). Our tests assume that Kyrgyz exhibits only one type or the other, but checking for a bimodal distribution in duration could be useful to discover if there is a contrast between the two types.

Kyrgyz does not normally allow even heterorganic vowels in hiatus; any given bound morpheme that may surface with an initial vowel never does so after another vowel, or alternatively creates a single long vowel. An example of the former pattern is the first person singular possessive suffix *-(I)m/*: after stems ending in consonants, it surfaces with a vowel, e.g. /tɨj-(I)m/ [tʰɨjim] ‘my tooth’, but after stems ending in vowels it surfaces without a vowel, e.g. /ɑłma-(I)m/ [ɑłmam] ‘my apple’. An example of the creation of a long vowel is /tɨfte-(U)U/ ‘bite-VN’, which surfaces as [tʰɨftø:]. Situations where hiatus of heterorganic vowels might occur are across word boundaries and in recent loan words. In the case of two short vowels across a word boundary within the same prosodic phrase, usually the second vowel deletes; e.g., /sarɯɑłma/ ‘yellow apple’ is normally realised as [sarɑłma]. Hiatus does seem to be possible if the first vowel is long, however: /damdu.ɑłma/ ‘delicious apple’ is normally realised as [damdu.ɑłma] (cf., Şabdan uulu & Batmanov 1933: 18), with shortening of the long vowel and an instance of hiatus. Additionally, deletion is much less prevalent across prosodic phrases, e.g. /ɟemɨjiɑłma/ ‘its fruit is an apple’ is usually pronounced [ɟemɨjiɑłma] with two immediately adjacent vowels. The reason deletion is not possible in this last example may also be, at least in part, due to morphological considerations: if /i/ is deleted, then no surface representation of the possessive morpheme remains. Şabdan uulu & Batmanov (1933: 18) provide similar examples that do witness deletion, e.g. /bałɑ ɯjɟadu/ ‘the child cried’ surfaces as [bałɯjɟadu].<sup>1</sup> Hiatus is also possible in many recent loanwords, e.g. [v.ɔdʰiˈtɔrʰɨjə] ‘lecture hall’, [v.ɨrɨˈpɔrt] ‘airport’, [ɨˈɔsʰɨf] ‘Joseph’, and [ˈxa.ɔs] ‘chaos’, where the first two vowels in each word are normally considered to be in separate syllables and to not form a diphthong, per Russian phonology. To our knowledge, none of these phenomena of hiatus in Kyrgyz have been studied phonetically.

In this study, we limit ourselves to exploring spectral dynamics (change in formants throughout the duration of vowel utterance) and F0 dynamics (change in pitch throughout the duration of vowel utterance) to investigate the question of dual-headed vowels. A change of spectral properties through the duration of a vowel would be indicative of a two-vowel analysis, since it would indicate that two distinct vowels are being produced—although it may not necessarily mean that the vowels were in distinct syllables. In terms of F0 dynamics, we believe that a change in pitch could occur between two vowels in hiatus—e.g., as the result of a glottal gesture—or across the span of two vowels in hiatus if the two syllables are stressed differently.

<sup>1</sup> A Kyrgyz-speaking consultant suggests that the first vowel is partially lengthened as well, as [baːɯjɟadu].

2.4. OTHER CONFOUNDS OF DURATION. Additionally, long and short vowels have been shown to have vastly different spectral properties in some languages (e.g., Hungarian, cf. Szende 1994). Kawahara (2016: 176–177) cites evidence of different—although much more similar—intrinsic spectral properties of Japanese short vowels and long vowels to argue that the long vowels cannot simply be two short vowels in hiatus.

It has also been shown that in some Turkic languages with contrastive vowel length, pitch may be a secondary cue of duration (e.g., Sakha, cf. Vasilyeva, Arnhold & Järvikivi 2016). Furthermore, it has been claimed that stress and vowel duration cannot be correlated in languages with contrastive vowel length—a falsifiable claim disputed by Lunden et al. (2017).

We aim to determine how Kyrgyz can contribute to these typologies. To do this, we examine spectral differences between long and short vowels, as well as the interaction of vowel duration with stress, syllable position in word, and other factors that often correlate with vowel duration, such as voicing of a following consonant, and presence in an open or closed syllable. A vowel is considered to be in a closed syllable if a following consonant is in the same syllable—i.e., if another consonant or word boundary follows the consonant immediately after the vowel.

**3. Methodology.** To assess the possibility of dual-headed long vowels in Kyrgyz, we examined spectral properties and F0 dynamics. Specifically, in the absence of extensive literature on dual-headed vowels, we assume that dual headed vowels ([V.V] sequences) could show either two spectral peaks (similar to a diphthong) or two different pitch peaks, while true long vowels ([V:] should show stable spectral dynamics and stable pitch dynamics. To assess what else duration might be correlated with, we identified potential confounds of duration, and attempted to account for the durational effects of each one by statistical means. Individual statistical measures are described with their results in section 4.

The corpus used (Washington 2016, 2019) consists of audio and ultrasound tongue imaging recordings of speakers of several Turkic languages. The stimuli were designed for examination of the articulation of vowels in each language—in the environment of a range of consonants, word lengths, and syllable structures—but studies to date using this corpus have ignored long vowels. The present study does not examine the ultrasound data, and examines data from only one of the five Kyrgyz speakers' data in this corpus.

The consultant examined in the present study (P03) was a female speaker of Kyrgyz from Naryn, Kyrgyzstan, at the time 31 years old, with proficiency in Russian and English and some knowledge of German. She was recorded in the Indiana University Speech Production Laboratory in 2015. Each target word form was embedded once in each of two carrier sentences: *Үйгө барып, \_\_\_\_\_ ден айттым* 'I went home and said \_\_\_\_\_' and *Үйгө кирип, \_\_\_\_\_ ден айттым* 'I entered the house and said \_\_\_\_\_'. Sentences were randomised and presented to the consultant on a screen six sentences at a time. More information on the stimuli, the procedure, and the participant(s) can be found in Washington (2016: particularly §4.2).

The stimuli read by the Kyrgyz-speaking consultant were injected into Praat (Boersma & Weenink 2019) TextGrid Tiers (for sentence, word, and vowel) using Python scripts. Each vowel token was delimited in the same TextGrid. The intervals were aligned by the authors and two additional students trained by the second author.

To remain agnostic in regards to the question of the phonological status of high unround [Vj(V)] sequences as discussed in section 2.2, these sequences were annotated according to their formant structure (and to some extent perceptual properties) and not on the basis of orthographic or morphological considerations. That is, when an instance of these sequences appeared to be [Vj(V)], it was annotated as such—i.e., with individual short vowels, differently from those which appeared to be [V:] or similar. Hence, in figures in this study, some [i] and [u] tokens are from /ij(i)/ and /uj(u)/ sequences where there was a clear articulation of [j], while all [i:] and [u:] tokens comprise exactly the set of those /ij(i)/

and /uj(w)/ sequences without an easily delimited [j].

In total, the analysed data comprised 2287 vowel tokens over 851 words. A series of Praat and R scripts were used to perform the measurements and analysis.

**4. Results.** This section reviews our contrast of long and short vowels along the following phonetic metrics: spectral properties (4.1), spectral and pitch dynamics (4.2), and the interaction between stress, length, and duration (4.3). It continues by comparing among short vowels (by far the larger part of our dataset) for a more detailed picture of confounds on duration, including but not limited to stress (4.4) and with more specific focus on syllable number (4.5).

4.1. SPECTRAL PROPERTIES. To get a sense of the extent to which short and long vowels have distinct spectral properties, we measured the first two formants (F1 and F2) at the midpoints of tokens of three short vowels and their corresponding long vowels.

It has been documented (cf. Hebert & Poppe 1963) that in Kyrgyz, palatal (or palato-alveolar) consonants have a huge impact on the quality of adjacent vowels. For this reason, measurements were compared only between vowels in analogous consonantal environments. Figure 1 shows aggregated data for three long/short vowel pairs when following a palatal consonant and preceding a non-palatal consonant. The particular vowels ([e], [e:], [a], [a:], [u], [u:]) were chosen to avoid visual clutter of overlapping in the vowel space, and to some extent based on the limitations of our data set.

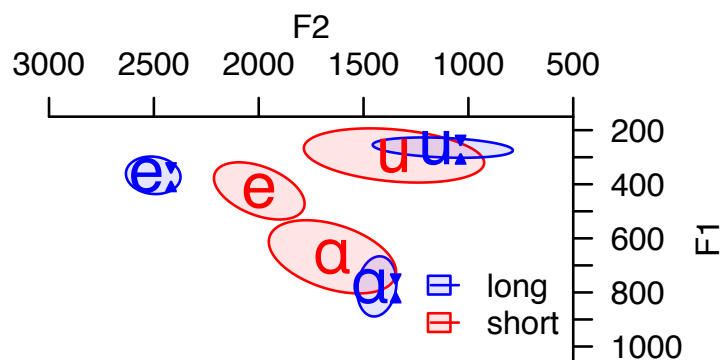


Figure 1: Midpoint formant values with standard deviation ellipses for three corresponding short (red) and long (blue) vowels, in the context of a preceding palatal consonant and a following non-palatal consonant. Formants were measured at the midpoint of vowels.

It can be seen that, in general, short vowels have more reduced (i.e., mid-central) formant values while long vowels are more peripheral. Also, long vowels appear more “precise”, with smaller standard deviation ellipses. Of the three short and long vowel pairs examined here, [e] and [e:] show the greatest difference, with standard deviation ellipses not overlapping at all, while the short and long members of the [a]~[a:] and [u]~[u:] pairs overlap considerably.

However, other factors may be contributing to the differences between these short and long pairs: in order to remove the confound of palatalization as discussed above, we were forced to select short and long vowels from somewhat different prosodic and consonantal contexts. The long vowels were all from the first syllable of one- to four-syllable words between /dʒ/ and /n/ (all were forms of the stems /dʒe:n/ ‘child of female relative’, /dʒa:n/ ‘rain’, and /dʒu:n/ ‘wash [oneself] up’). The short vowels were all in the second syllable of two- to four-syllable words, following /ʃ/ (except for [a]), /tʃ/ (except for [u]), and /j/, and preceding a range of consonants, including /m/, /p/, and /n/ (for [u]), and /r/ and /t/ (for [e] and [a]). These differences between the contexts for the short and long vowels are in part a limitation of the dataset examined, but are also due to some extent to limitations of the Kyrgyz lexicon—although a more

balanced dataset is indeed possible.

While this section presents spectral differences between short and long vowels akin to those documented in other languages, it may be that the differences seen are in fact related to prosodic or consonantal effects.

4.2. SPECTRAL AND PITCH DYNAMICS. The quantifiable acoustic properties that would be expected to differentiate long vowels and heterosyllabic short vowels have not been well established in the literature. We posit that a sequence of heterosyllabic short vowels may have two peaks in F0 or diphthong-like spectral properties (changing over time), whereas phonetically long vowels would have level F0 and stable spectral properties. Time was normalised to the annotated “edges” of the vowels so that long and short vowels could be effectively compared.

To consider which implementation the Kyrgyz long vowels exhibit, we measured the spectral dynamics of each vowel throughout its duration, quantified as F2-F1, similar to the quantification of spectral dynamics used by Mielke, Carignan & Thomas (2017) (Figure 2), as well as the dynamics of F0 throughout each vowel’s duration (Figure 3). Vowels were controlled for palatalization effects from their consonantal environment, as in section 4.1.

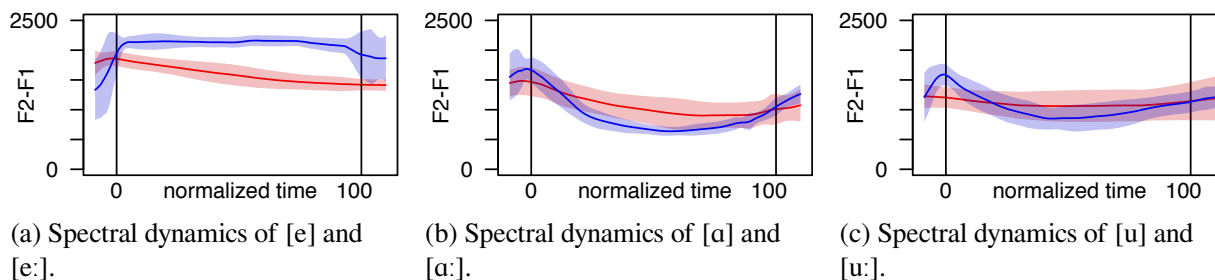


Figure 2: Spectral dynamics (F2-F1) of three long (blue) and short (red) vowel contrasts through the duration of the vowel (0%-100%) as produced by participant P03, after palatal consonants and before non-palatal consonants. Lines represent mean and bands represent one standard deviation.

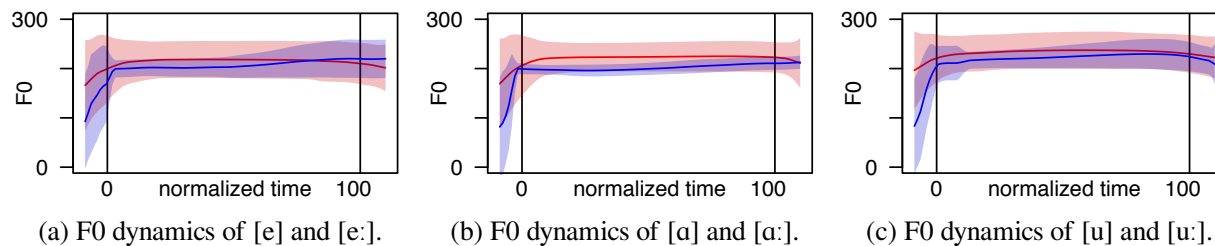


Figure 3: Dynamics of F0 throughout three long (blue) & short (red) vowels, produced by participant P03. Lines represent mean and bands represent one standard deviation.

The spectral dynamics of long and short vowels are distinct for some short/long pairs. In [a]/[a:] and [u]/[u:] pairs, we see evidence of what appears to be undershoot in the short vowels—there is not as great a “dip” after the palatal consonant, meaning the tongue did not reach as extreme a position. The remaining differences in spectral dynamics are thought to mostly be due to the different consonantal contexts, as discussed in section 4.1. However, aside from the effect from preceding palatals (the initial dip downwards for most of the vowel types), there appears to be no evidence of diphthongal qualities, suggesting that these are straightforwardly long vowels.

The F0 dynamics are also almost entirely flat, do not have multiple peaks, and do not appear drastically

different for short and long vowels. The somewhat lower apparent F0 of long vowels could be related to the difference in syllable position.

Neither of the measures we used—spectral dynamics and pitch dynamics—support an analysis of long vowels as dual-headed in Kyrgyz.

4.3. STRESS AND DURATION. The distribution of durations of the examined Kyrgyz vowels by length and stressed/unstressed status are shown in Figure 4. As determined by t-tests, a significant difference in duration was observed between short and long vowels overall, as well as between each short vowel and its corresponding long vowel, disregarding stress. This corroborates that duration is indeed a large piece of the vowel length distinction in Kyrgyz. Regarding stressed vs unstressed vowels, stress and duration are significantly correlated for some vowel pairs, though syllable position may be the ultimate cause of this correlation (see section 4.4). It was also found that short high vowels ([i], [y], [u], [ɯ]) are overall significantly shorter than short non-high vowels ([e], [ø], [ɑ], [o]), as reported in Washington (2016). This difference has not yet been explained in the literature. As shown in section 4.5, however, this duration difference is robust to other duration-altering effects.

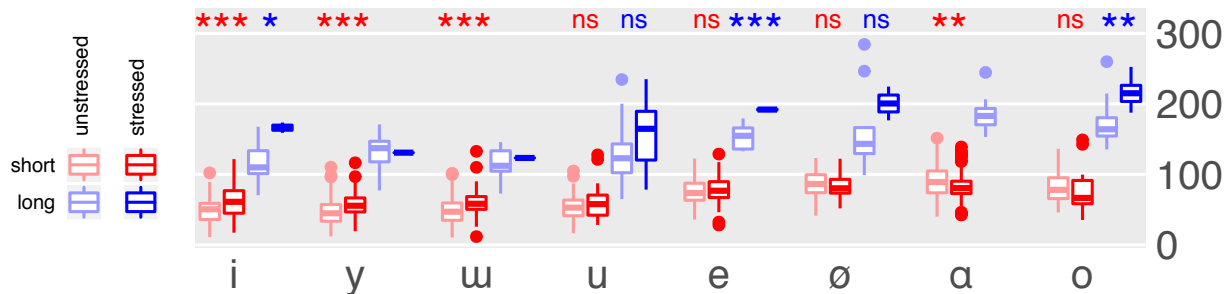


Figure 4: Durations (ms), by short (red) and long (blue) vowel and unstressed (lighter) / stressed (darker). Stars along the top indicate level of significance in duration between stressed and unstressed reflexes of each vowel, as determined by t-tests. \*\*\*:  $p < 0.001$ , \*\*:  $p < 0.01$ , \*:  $p < 0.05$ , and ns: not significant.

4.4. CONFOUNDS OF DURATION. In order to better understand the interaction between stress and duration, we investigated the effects of four potential confounds noted cross-linguistically as correlating with duration: stress (or lack thereof), location within the word of the syllable that contains the vowel, openness of syllable (whether the syllable included a coda or not), and voicing of the consonant following the vowel. Only short vowels were analysed in these environments ( $n=1587$ ). A linear regression including all four variables was run for each vowel, the results of which are presented in Table 1.

Syllable location was divided into initial, medial, and final syllables, and was dummy coded, with the medial syllable chosen as the reference level. (The medial syllable can't be compared to itself, hence the 'N/A's in that row.) To prevent tokens from being analysed twice, vowels in monosyllabic words were excluded from the analysis. Various models using syllable number instead were also tried, and found to be less predictive of our data than this model. These included dividing syllables by number, both counting from the beginning of a word (initial = 1, peninitial = 2, etc.) and counting from the end of a word (final = 1, penultimate = 2, etc.).

All words sampled had final stress (the default in Kyrgyz), so stress as a predictor variable was collinear with syllable location across all vowels. For [o], voicing of the following consonant was also collinear with syllable position: this is likely due to a quirk of vowel and consonant distributions in this dataset. While for all other vowels, between 50 and 75% of their tokens were followed by voiced consonants, [o] was followed by a voiced consonant 85% of the time. This is not a known limitation of the Kyrgyz lexicon, and can be remedied in future studies.

	i	y	ʊ	u	e	ø	ɑ	o
initial syllable	*		***					***
medial syllable	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
final syllable	***	***	***		**		**	***
—C <sub>[+voice]</sub>			*					coll.
open syllable				*			**	
stressed	coll.	coll.	coll.	coll.	coll.	coll.	coll.	coll.
Adjusted R <sup>2</sup>	0.1603	0.111	0.08781	0.04377	0.1749	0.05076	0.1128	0.1187
p-value	4.083e-05	6.505e-05	1.78e-05	0.03873	1.479e-07	0.02323	3.276e-09	0.0006457
	***	***	***	*	***	*	***	***

Table 1: Linear regressions of variables by vowel, with syllable locations (compared to medial syllable) as well as other potential confounding factors. “N/A” is used to indicate invalid comparisons, and “coll.” stands for collinear—in this case with syllable position. Stars represent the predictive power of each variable for each vowel, according to its p-value.

For the potential confounds unrelated to syllable position—syllable openness and the voicing of the following consonant—effects were observed only in isolated vowels. Because of our relatively small sample size, and the fact that eight regressions were run (one per vowel), we conclude that these two conditioning environments do not broadly affect vowel duration in Kyrgyz, and may in reality have no significant effect.

For syllable position, a vowel’s presence in the final (stressed) syllable was observed to significantly affect its duration in all models where  $p < 0.01$ : the models for /i/, /y/, /ʊ/, /e/, /ɑ/, and /o/. These are also the models that predict the largest percentages of the data (R<sup>2</sup>). This is a marked change from Figure 4, where stress alone was an inadequate predictor of duration for the majority of vowels, both long and short. Though ‘presence in the initial syllable’ did not have widespread predictive power across vowels in Table 1, the separation of initial and medial syllables in this model is likely responsible for the effect applying to more final/stressed vowels as compared to Figure 4.

There may, in fact, be more going on with the duration of vowels in word-initial syllables; this will be discussed further in section 4.5.

4.5. SYLLABLE POSITION. Although a vowel in a final syllable can be predicted with confidence to be different in duration from a vowel in a word-medial syllable in Kyrgyz for most vowels, as shown in Table 1, the direction of this effect (i.e., lengthening or shortening) depends on the height of the vowel. Figure 5 shows a graph of duration by syllable position for each vowel, and Table 2 shows the significant differences in duration by syllable position for each vowel, calculated by Tukey’s range test.

As in section 4.4, vowels in word-final syllables in particular are significantly different in duration from vowels in medial syllables across most vowel phonemes. Initial-syllable vowel duration as compared to vowel duration in other positions, however, is more distinct in this analysis. Firstly, different initial-syllable vowels are calculated to be significantly different in duration from medial-syllable vowels—/i/, /y/, /e/, and /ɑ/ in Table 2, as opposed to /i/, /ʊ/ and /o/ in Table 1. Additionally, some significant differences in duration between initial- and final-syllable vowels appeared (for /ʊ/ and /o/), which were not able to be calculated in Table 1.<sup>2</sup>

The differing analyses of initial syllables in no way contradict the analysis that final syllable stress is related to a significant duration difference across vowels. The more ambiguous data related to word-initial syllable duration may help illuminate a less understood prosodic phenomenon in word-initial syl-

<sup>2</sup> We note that [u] was the only vowel for which there were no significant findings in these tests. In no test in the entire study did [u] reach the significance threshold. This may be due to limitations in the dataset.



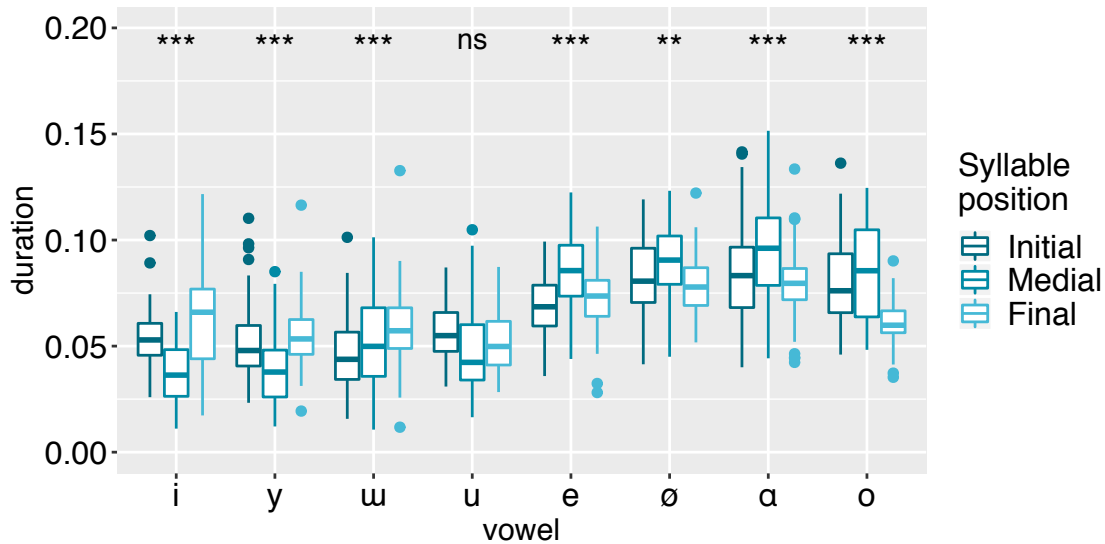


Figure 5: Duration by syllable position for short vowels. Stars for each vowel are the significance that not all means are the same, from an ANOVA test.

i		y		u		u		e		ø		a		o	
1	-1	1	-1	-1		ns		m		m		m		1	m
m		m		1	m			1	-1	-1		1	-1	-1	

Table 2: Summary of Tukey HSD results. Vowels in higher-placed syllables were determined to be significantly longer than vowels in lower-placed syllables. 1 = vowel in word-initial syllable, m = vowel in word-medial syllable, -1 = vowel in word-final syllable.

lables in Kyrgyz: “word-initial prominence.” In many Turkic languages, a syllable other than the final one (often the first syllable of long native words like the ones used in this study) receives some demarcation of accent beyond the normally word-final primary accent. Some sources discuss this secondary accent as a “stress accent” in contrast to the word-final “pitch accent”, cf. Johanson (2006: 34-35) for Turkic generally and Kirchner (2006: 320) for Kazakh. Imart (1981: §683, §731) summarises other sources’ discussions of a secondary accent in Kyrgyz, and, noting the mobility of this accent for purposes of focus and emphasis, argues (*ibid.*: §738-§739) that this is not an accentual phenomenon at all, but is the realisation of pragmatic intonation. Alternatively—and perhaps not at odds with other analyses—this initial syllable prominence could be exactly that: an instantiation of initial-syllable prominence (cf., Barnes 2003), also examined as domain-initial strengthening—both articulatory (e.g., Cho & Keating 2001) and acoustic/perceptual (e.g., Cho, McQueen & Cox 2007)—and contextualised by the general literature on phonetic and articulatory strengthening at [especially prosodic] boundaries (e.g., Byrd & Saltzman 2003). Fougeron & Keating (1997), for example, find that vowels are articulatorily stronger (as measured by less linguo-palatal contact) in domain-final positions.

Because the conditioning environments and semantic or pragmatic cues for initial-syllable prominence are not yet known, we cannot definitively identify its presence in our data, despite potentially observing its effects. Positing a phenomenon of this sort may explain why only some initial vowels pattern with (stressed) final vowels, and why the linear model and Tukey’s range test predict that different vowels exhibit significantly different duration in initial syllables relative to other syllables—e.g., certain initial

syllables may be receiving this prominence more than others, and/or the prominence may be cued by some phonetic environments that are unequally distributed in our data.

We can think of at least two potential explanations for the observed differences in vowel duration between edge- and medial-syllable positions across vowel heights. Under one hypothesis, high vowels are intrinsically much shorter than non-high vowels in Kyrgyz, as seen in word-medial syllables. Then this durational distinction mostly neutralises in word-edge syllables. Under another hypothesis, due to some other neutralisation process between high and non-high vowels in medial position, duration is exaggerated in medial syllables to distinguish between high and non-high vowels in that context. To choose between these hypotheses, it would be necessary to examine spectral properties of the vowels by syllable position. If the formant values of either high and non-high vowels are more similar to one another in word-medial syllables, then that would suggest that spectral neutralisation is occurring in word-medial syllables, supporting the second hypothesis. A lack of spectral neutralisation by position would support the first hypothesis. However, a third possibility is that a spectral distinction by syllable position (if one were found) could be in part an exaggeration to make up for a mostly neutralised duration distinction. The limits of our dataset prevent us from reliably deciding between these hypotheses. McCollum's (2019) findings that non-initial vowels are less peripheral than initial vowels in Kyrgyz may be relevant, in that this centralisation effect may be the phenomenon that duration in word-medial syllables is sensitive to. A combined study of spectral and durational properties of Kyrgyz vowels is called for. This study is needed in particular because both hypotheses we offer are incompatible with word-edge strengthening as a cause for the observed effects. Specifically, the first hypothesis is that Kyrgyz exhibits word-edge *neutralisation*, while the second hypothesis is that Kyrgyz exhibits *word-medial* strengthening. Because of this, we believe that the data examined in this study does not lead to an analysis of word-edge strengthening, and that another mechanism is at play. This perhaps also calls into question previous notions of how the already poorly understood accentual system of Kyrgyz works.

**5. Conclusions.** Finding a strong correlation between length and duration, no evidence that all vowels in Kyrgyz are dual-headed, no uniform diphthongal tendencies, and only weak evidence for the correlation of spectral factors with the length distinction, we suggest that Kyrgyz long vowels differ from short vowels mainly in duration.

Despite finding no evidence of widespread dual-headed vowels, we have noted some vowels in our corpus which do seem dual-headed, both perceptually to us, and in some of their acoustic properties. An example of this is presented in Figure 6, which shows data from the first vowel in the word [q<sup>h</sup>u:sunbu] /qu:-sIn-bI/ 'chase-IMP.3-QST' "should he/she/it/they chase?", as read by another Kyrgyz-speaking consultant, P04.

We note two acoustic properties that may contribute to the categorisation of this long vowel as dual-headed. First of all, it appears to have an intensity peak in its first third, a subsequent dip in intensity in the middle third, and then a bit of a rise again in intensity before the vowel ends and the following consonant begins. Additionally, there is a steady rise in pitch through the vowel. In the words we examined from this speaker, higher pitch was often noted on stressed syllables. In most words of this form, stress would be on the syllable before the optative morpheme, e.g. [t<sup>h</sup>iʃ'tesinbi] /tiʃte-sIn-bI/ 'bite-IMP.3-QST' "should he/she/it/they eat?". The rise in pitch in the second half of this long vowel may indicate that it is acting like two individual vowels, [u.u], where the first is unstressed and the second is stressed. It is also worth noting that we saw slight pitch rises—as well as larger standard deviations in F0—in Figure 3 for two of the three long vowels whose F0 dynamics we investigated. The larger standard deviation may suggest that some of the vowel tokens exhibited a pitch rise, while others did not. Based on these preliminary observations, future work should investigate whether intensity could be a sign of dual-headed vowels, and also take into account that there may be a bimodal distribution—i.e., that some long vowels may be two individual vowels /V.V/, while others are simple long monophthongs /V:/. It may also be

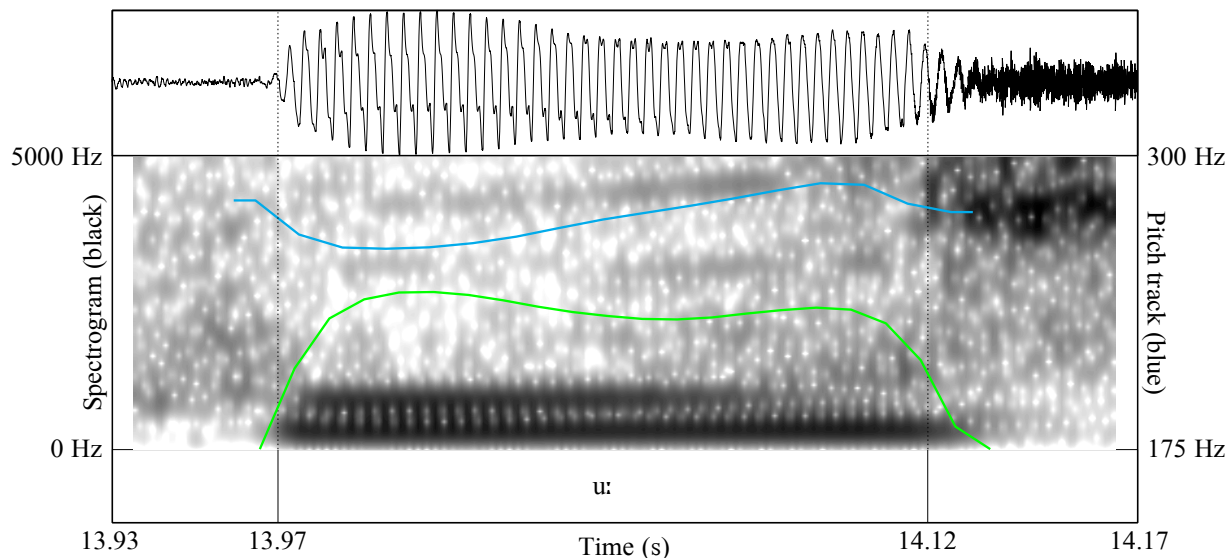


Figure 6: An example of an apparent dual headed vowel from our corpus. The vowel is [u:] from the word [q<sup>h</sup>u:sunbu] as uttered by P04. A pitch track is in blue and a measure of intensity is in green. Created in Praat (Boersma & Weenink 2019).

fruitful to compare long vowels to cases of cross-word hiatus of homorganic short vowels where deletion does not occur.

More generally, we plan to analyse data from more speakers to determine whether our findings hold more broadly across Kyrgyz speakers.

In this study, we demonstrated that Kyrgyz short vowels can be divided into four measurably distinct durational categories by vowel height and syllable position: (in ascending order of duration) high/word-medial, high/word-edge, non-high/word-edge, non-high/word-medial. We offer two competing hypotheses for a mechanism that could produce these categories, and a potential methodology to distinguish between them. In finding that initial and final syllables pattern together, we contribute to the understanding of accent and edge effects in Kyrgyz, and Turkic languages as a whole. However, we do not believe that we have identified a case of edge-position strengthening.

In conducting this study, we have also contributed to the understanding of long-standing questions concerning the epistemology of long high unrounded vowels in Kyrgyz, as well as to the debate about whether duration can be a correlate of stress in a language with contrastive vowel length. In future work, besides examining data from more Kyrgyz speakers, we plan to analyse the ultrasound data corresponding to the audio data examined here in order to more thoroughly address these questions, and to test the competing hypotheses as to whether or not there are—aside from duration—significant articulatory differences (i.e., in tongue position and configuration) between long and short vowels in Kyrgyz.

## References

- Barnes, Jonathan. 2003. Initial-syllable prominence: what is it and where does it come from? Handout. <http://www.bu.edu/linguistics/UG/barnes/Syll%201%20MIT%2003.pdf>.
- Boersma, Paul & David Weenink. 2019. Praat: doing phonetics by computer. Version 6. <http://www.praat.org/>.

- Byrd, Dani & Elliot Saltzman. 2003. The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31. 149–180. doi: [https://doi.org/10.1016/S0095-4470\(02\)00085-2](https://doi.org/10.1016/S0095-4470(02)00085-2).
- Cho, Taehong & Patricia A. Keating. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29. 155–190. doi: <https://doi.org/10.1006/jpho.2001.0131>.
- Cho, Taehong, James M. McQueen & Ethan A. Cox. 2007. Prosodically driven phonetic detail in speech processing: the case of domain-initial strengthening in English. *Journal of Phonetics* 35. 210–243. doi: <https://doi.org/10.1016/j.wocn.2006.03.003>.
- Fougeron, Cécile & Patricia A. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101(6). 3728–3740. doi: <https://doi.org/10.1121/1.418332>.
- Hebert, Raymond J. & Nicholas Poppe. 1963. *Kirghiz manual*. Vol. 33 (Uralic and Altaic Series). Bloomington: Indiana University.
- Imart, Guy. 1981. *Le kirghiz (Turk d'Asie Centrale Soviétique): Description d'une langue de littérisation récente*. Aix-en-Provence: Publications de l'Université de Provence.
- Johanson, Lars. 2006. The structure of Turkic. In Lars Johanson & Éva Ágnes Csató (eds.), *The Turkic Languages*, chap. 3, 30–66. Routledge.
- Kawahara, Shigeto. 2016. Japanese has syllables: a reply to Labrune. *Phonology* 33(1). 169–194. doi: <https://doi.org/10.1017/S0952675716000063>.
- Kirchner, Mark. 2006. Kazakh and Karakalpak. In Lars Johanson & Éva Ágnes Csató (eds.), *The Turkic Languages*, chap. 19, 318–332. Routledge.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
- Lunden, Anya, Jessica Campbell, Mark Hutchens & Nick Kalivoda. 2017. Vowel-length contrast and phonetic cues to stress: an investigation of their relation. *Phonology* 34(3). 565–580. doi: <https://doi.org/10.1017/S0952675717000288>.
- McCollum, Adam. 2019. *Gradience and locality in phonology: case studies from turkic vowel harmony*. UC San Diego Doctoral Dissertation. <https://escholarship.org/uc/item/7sx31303>.
- Mielke, Jeff, Christopher Carignan & Erik R. Thomas. 2017. The articulatory dynamics of pre-velar and pre-nasal /æ/-raising in English: an ultrasound study. *The Journal of the Acoustical Society of America* 142(1). 332–349. doi: <https://doi.org/10.1121/1.4991348>.
- Roberts-Kohno, R. Ruth. 1995. Vowel coalescence and hiatus in Kikamba. In Akinbiyi Akinlabi (ed.), *Theoretical approaches to African linguistics*, vol. 1 (Trends in African Linguistics), 313–327. Trenton, New Jersey: Africa World Press.
- Şavdan uulu, Aşman & I. A. Batmanov. 1933. *Qыроғыз тилинiн ментардығ оғраматикеси*. Prunza, Taşkent: Qығметмбас, Orasmembas.
- Szende, Tamás. 1994. Hungarian. *Journal of the International Phonetic Association* 24(2). 91–94. <http://www.jstor.org/stable/44526167>.
- Vasilyeva, Lena, Anja Arnhold & Juhani Järvikivi. 2016. Phonetic correlates of phonological vowel quantity in yakut read and spontaneous speech. *The Journal of the Acoustical Society of America* 139. 2541–2550. doi: <https://doi.org/10.1121/1.4948448>.
- Washington, Jonathan. 2016. *An investigation of vowel anteriority in three Turkic languages using ultrasound tongue imaging*. Indiana University Doctoral Dissertation. <http://hdl.handle.net/2022/20954>.
- Washington, Jonathan. 2019. An investigation of the articulatory correlates of vowel anteriority in Kazakh, Kyrgyz, and Turkish using ultrasound tongue imaging. *University of Pennsylvania Working Papers in Linguistics* 25. <https://repository.upenn.edu/pwpl/vol25/iss1/24/>.
- Батманов, И. А. 1938. *Северные диалекты киргизского языка*. Фрунзе: Киргизгосиздат.
- Батманов, И. А. 1939. *Грамматика киргизского языка*. Vol. 1. Фрунзе, Казань: Киргизгосиздат.

Уртегешев, Н. С. 2017. Шорский язык: сложный или простой? *Языки и фольклор коренных народов Сибири* 34. 4–15.