Undershoot in Kyrgyz short vowels is articulatorily conditioned

Nathaniel Ziv Stern[1] *and Jonathan Washington*[1]

[1]Swarthmore College (USA)

nzivstern@gmail.com, jonathan.washington@swarthmore.edu

Introduction: This study investigates the question of whether previously identified spectral differences between Kyrgyz short (V̆) and long (Vː) vowels (Vs) (Stern & Washington 2019) are due to articulatory undershoot. Languages with small spectral differences between V lengths (e.g. Creek, Johnson & Martin 2001; and Japanese, Yazawa & Kondo 2019) have been claimed to exhibit centralisation of the V̆s due to undershoot (Lindblom 1963), though the degree to which this is caused by timing and/or adjacent segments is debated (van Son 1993). The present study finds that following coronal consonants lead to undershoot in Kyrgyz back V̆s, but not in Vːs or front Vs.

Methodology: The corpus used consists of audio and ultrasound (US) tongue imaging recordings of speakers of several Turkic languages. The stimuli were designed for examination of the articulation of Vs in the environment of a range of Cs, word lengths, and syllable structures. More information, including about participant(s), in Washington (2016, particularly §4.2).
This study examines recordings of a 42-year-old cis-female Kyrgyz speaker from Suzaq district, Jalalabat oblast, Kyrgyzstan with proficiency in Russian, Turkish, English, and some Arabic. We examine half the V inventory of Kyrgyz: the 8 rounded V phonemes, grouped as 4 pairs that contrast only in length: [y]/[yː], [ø]/[øː], [o]/[oː], [u]/[uː]. Vs were measured in the environment following dorsal stops (K: [k]/[q]) and preceding coronal obstruents (D) in the first syllable of polysyllabic words. This environment and the particular set of Vs were chosen to avoid effects of stress and consonantal influence on V acoustics, and due to limitations of the corpus.
The tongue surface was manually traced and US and audio were aligned using UltraTrace (Murphy et al. 2020). Formants were measured at the midpoint of each annotated V token. Tongue traces of the US frame closest to the V midpoint were used (US frames every 19.6ms). In total, the analysed data comprised 162 V tokens over 2 repetitions each of 81 unique word forms.

Findings: Figure 1 presents formant measurements for the 8 V categories examined, showing sightly more peripheral measurements of the Vːs, but only back ones. Figure 2 shows tongue traces for the 4 V̆/Vː pairs. The main difference in the tongue position of the pairs is that the back V̆s have higher tongue tip position than back Vːs. V̆s exhibit some properties of the tongue position of the preceding K, most noticeable among high Vs.
Our interpretation of the cause of the former difference is that in V̆s, the tongue tip is already partially raised for the articulation of the following D, while in Vːs the articulatory V target is more fully reached, with a lower tongue tip—a form of articulatory undershoot occurring due to the short span of time from the articulatory peak of a V̆ to the following C. We posit that front V̆s are not subject to the same form of undershoot because the front of the tongue is already raised to articulate the front V. We attribute the less peripheral spectral properties (relative to Vːs) of back V̆s but not front V̆s to this difference in tongue tip position. This explains the greater acoustic and articulatory similarity between front V̆s and front Vːs.

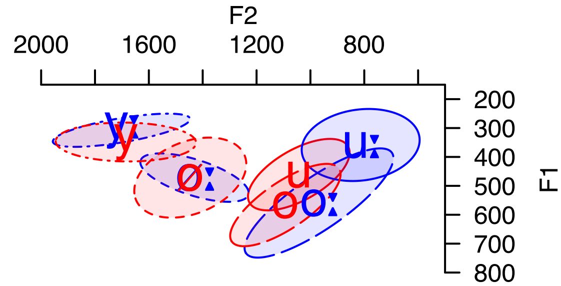

Figure 1: Standard deviation ellipses around the midpoint for each of the 8 V categories examined (short V red, long V blue).

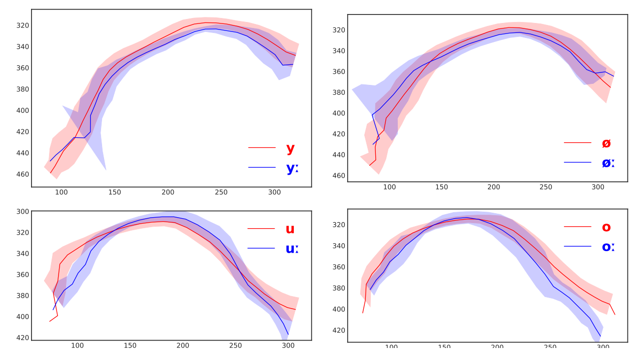

Figure 2: Tongue trace averages (polar coordinates) for each V category, with standard deviation bands. Tongue tip to right.

Future work: To test the undershoot hypothesis, we plan to compare the articulation of adjacent Cs to the Vs, and model the expected acoustic effects of the noted articulatory differences. We also plan to measure equivalent stimuli for other speakers of Kyrgyz in the corpus, and explore differences between Kyrgyz V̆s and Vːs in other articulatory contexts. We would additionally like to examine the extent to which undershoot effects that appear in the presence of a V length contrast are dependent on V and place of articulation of adjacent Cs cross-linguistically, or whether this type of effect can be language-dependent.

References

Johnson, K., & Martin, J. (2001). Acoustic vowel reduction in Creek: Effects of distinctive length and position in the word. *Phonetica*, *58*, 81–102. https://doi.org/10.1159/000028489

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, *35*(11), 1773–1781.

Murphy, K., Stern, N. Z., Swanson, D., Ho, C., & Washington, J. (2020). UltraTrace: A free/open-source cross-platform tool for manual annotation of ultrasound tongue imaging data. *UltraFest IX*.

Stern, N. Z., & Washington, J. (2019). A phonetic study of length and duration in Kyrgyz vowels. *Proceedings of the 4th Workshop on Turkic & languages in contact with Turkic (Tu+4)*. https://doi.org/10.3765/ptu.v4i1.4577

van Son, R. (1993). *Spectro-temporal features of vowel segments* (Doctoral dissertation). Universiteit van Amsterdam.

Washington, J. (2016). *An investigation of vowel anteriority in three Turkic languages using ultrasound tongue imaging* (doctoral dissertation). Indiana University. http://hdl.handle.net/2022/20954

Yazawa, K., & Kondo, M. (2019). Acoustic characteristics of Japanese short and long vowels: Formant displacement effect revisited.

Keywords: vowel length, Interaction of phonology & phonetics, Kyrgyz