Velarization of Russian labial consonants

Kevin D. Roon[1,2] and D. H. Whalen[1,2,3]

[1] CUNY Graduate Center, [2] Haskins Laboratories, [3] Yale University
kroon@gc.cuny.edu, dwhalen@gc.cuny.edu

ABSTRACT

Contrast between palatalized and non-palatalized consonants is a robust feature of Russian. Non-palatalized consonants are generally viewed as velarized (with secondary dorsal articulation) rather than "plain" (without secondary articulation). This velarization has been proposed to play an important role in the phonology of the language. The bases for velarization are transcriptions and acoustics, as well as physiological data from liquids. There is little systematic articulatory investigation of whether consonants other than liquids are velarized. The present study used ultrasound data to determine whether there were identifiable dorsal articulations in labial stops, fricatives, and nasals, both word-initially and word-finally. We looked at labials to avoid confounding effects of other lingual targets. We found that non-palatalized Russian labial consonants have a velar constriction associated with them, irrespective of manner. We further found that the tongue-dorsum gesture associated with labial consonants had a notable coarticulatory effect on the neighboring vowel.

Keywords: Russian, velarization, ultrasound

1. INTRODUCTION

Russian has a robust contrast between palatalized and non-palatalized consonants across manner, primary oral articulator, voicing, and word/syllable position, in both stressed and unstressed syllables [11, 14, 16, 21, 29]. It has been argued that Russian non-palatalized consonants are in fact velarized [8, 9, 20, 21, 26, 30], i.e., they include a dorsal, velar gesture. There are differing claims as to the exact nature of this dorsal gesture, with some [21, 26] describing it as a full velar gesture similar to that associated with [ɯ], others [1, 17] describing it as a less extreme articulation, and still others [4, 11] claiming the retraction noted in "velarization" is a reflex of a pharyngealization.

Another claim is that the target [18] or presence [17] of velarization depends on the particular segment or on the primary oral articulator [15]. It has also been claimed that secondary velarization is optional [16], though whether this optionality is within- or across-speaker is not stated explicitly.

Most of the experimental work on velarization has relied on acoustic data [8, 9, 20, 21]. Investigations using articulatory data have been limited to qualitative assessments of a small amount of data [4, 26] or few speakers [15]. An exception is a study by Litvin [18], which used ultrasound data [27] from 6 speakers. Results from [18] contradicted the claims that velarization is a reflex of pharyngealization and that [s] lacks velarization since it is coronal [15]. Litvin found that whether the dorsal gesture was velar or uvular was segment-specific, and also concluded that the dorsal gesture was inherent to the consonant since it was unaffected by the following vowel. Litvin looked only at [l] and the fricatives [f, s, ʂ, x], all word-initially. Since the realization of palatalization varies based on word position and manner [8], the realization of velarization may also depend on these factors.

Velarization of non-palatalized consonants has been argued to play an important role in the phonology of Russian [21], as in other languages [2]. The present study was therefore designed to add to the findings from [18] by using ultrasound imaging of the tongue to determine whether there are identifiable dorsal gestures associated with non-palatalized labial consonants in Russian, across three different manners and in two different word positions.

2. METHODS

The experiment was conducted in the Speech Production, Acoustics, and Perception Laboratory at the CUNY Graduate Center. The experiment was approved by the CUNY Institutional Review Board.

2.1. Participants

Three native speakers of Russian, all originally from Moscow and currently living in the New York City metro area, took part in the experiment: one 29-y.o. female, one 28-y.o. male, and one 32-y.o. male. All speakers reported that they had no speech, language or hearing disorders, and provided informed consent.

2.2. Procedure

2.2.1. Stimuli

Stimuli consisted of 6 CVC syllables, shown in Table 1. The segments of interest in all stimuli were labial consonants: the stop [p], the fricative [f], and the nasal [m], which appeared both word-initially and word-finally. Labial consonants were chosen since the primary oral articulator does not involve the tongue, so any observed regular gestures of the tongue would be unlikely to be attributable to the influence of some other lingual target. It has also been claimed [17] that the dorsal gestures associated with labials are less extreme than those of the liquid [ɫ], so any positive evidence for velarization of labial consonants would be more compelling. To control for voicing, only voiceless obstruents [p, f] were used since word position was manipulated in the experiment and voiced obstruents devoice word-finally in Russian [11], and thus voiced obstruents could not be compared across word positions. V was always [a].

Table 1: Stimuli. Real words in Russian are denoted with an asterisk. Others are nonce words.

	Word-initial	Word-final
Stop	pam	map
Fricative	fam	maf
Nasal	mat*	tam*

Stimuli included real and nonce words, since it was not possible to construct CVC stimuli with only one or the other. Each Russian speaker produced all of the target stimuli in the carrier phrase [a ɛtə ____] (‘and this is a ____ ’) presented on a computer screen in Russian orthography. Each stimulus appeared ten times, in randomized order. These stimuli were interspersed among a larger set of CVC stimuli that were collected for other experiments.

2.2.2. Data collection

Ultrasound images of the lingual articulation along the midsagittal plane were recorded using an Ultrasonix SonixTouch ultrasound machine. Speakers sat in a chair and placed their chin on the ultrasound transducer (C9-5/10 micro-convex), which was held by a custom-made metal arm. The transducer holder contained a spring that allowed for movement of the jaw during speech. The display of the ultrasound machine was streamed to a PC and captured at 59.9402 Hz with an video capture card using a lossless codec (Magic YUV). An Optotrak Certus system was used to track the position of transducer holder and each speaker’s head so that the data could be corrected in post-processing for movement of the speaker’s head relative to the transducer. The movement of the transducer holder was tracked with infrared markers, and each speaker’s head movement was indexed using four infrared markers: one at the nasion and three on the forehead [24]. Audio was captured using a Sennheiser shotgun microphone, concurrently recorded both by the Optotrak system at 44.1 kHz and by the video capture card at 48 kHz. Optotrak data and ultrasound videos were synchronized based on the acoustics.

2.2.3. Data quantification and analysis

The following were identified and labelled by hand based on the acoustic record using Praat [3]: the durations of each labial of interest (Table 1) and of the initial [a]s of the carrier phrase from the first ten tokens, and the release of [p]s. The temporal midpoint of each segment (except [p]) was then calculated.

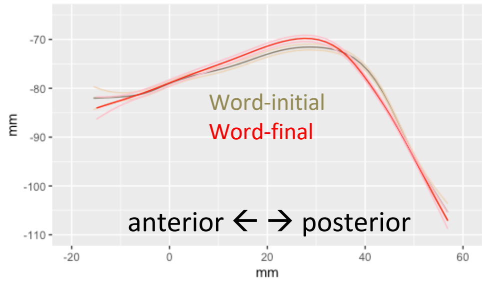
Video frames corresponding to the midpoints and releases above were identified, and the surface of the tongue was traced by hand, using GetContours [28], with each contour being defined as a series of 100 *xy* pixel values. Contours were converted to mm and head-corrected following [31]. While the contours were all defined in relation to a fixed head reference, the coordinate system shown in the figures below is arbitrary and does not have an origin corresponding to any fixed anatomical point in the vocal tract. Problems with recording the Optotrak resulted in the loss of 10 tokens for speaker 2 (1 token each of [fam] and [tam], 2 tokens of [pam], and 3 tokens of [mat] and [map]). One speaker produced 11 tokens of [pam] and [tam], and another produced 11 tokens of [tam].

Smoothed cubic splines defined in Cartesian coordinates were calculated for given groups of contours along with 95% Bayesian confidence intervals for all groups using the gss package [10] in R [23]. Areas where the confidence intervals did not overlap for at least 1 cm were interpreted as significantly different [6]. It has been argued [19] that polar coordinates are more appropriate than Cartesian coordinates for calculating smoothed splines for tongue contours. However, polar coordinates are appropriate when a virtual probe origin can be identified [12]. The identification of such a virtual origin is possible when the ultrasound probe is stabilized relative to the participant’s head, e.g., [25]. Since the present method relied on post-hoc head correction of the contours rather than head stabilization, the identification of a single origin corresponding to the probe was not appropriate, and thus Cartesian coordinates were used.

3. RESULTS

The first comparison was of the tongue shapes of the labials within word position, for a total of 18 comparisons (3 segments x 2 word positions x 3 speakers). There were only six significant differences between any two segments: a section of approximately 1 cm the middle of the tongue contours

Figure 1: Cubic spline fits of tongue contours with 95% Bayesian confidence intervals for word-initial (gold) vs. word-final (red) labials [p, f, m], for speaker 2.



of [f] were significantly higher than those of [p, m] for speaker 3, word-final [m] was significantly lower than [p, f] for speaker 2, and word-final [m] was significantly anterior to [p, f] for speaker 3.

Figure 1 shows the splines fit for contours from all three segments [p, f, m] by word position for speaker 2, whose pattern is indicative of all speakers (plots for other speakers omitted due to space considerations). There is a noticeable upward arching of the tongue dorsum (i.e., roughly the mid third of the contour) for all three speakers, irrespective of word-position, which we investigate further below. For all three speakers, the posterior part of the tongue is more advanced for word-final labials than for word-initial labials. In addition, the tongue dorsum was higher for speakers 1 and 2 word-finally (with a similar trend for speaker 3). The combination of these differences indicates that if this tongue configuration reflects a velarization gesture associated with the labials, it has a greater extent word-finally than word-initially.

In order to establish whether the observed tongue-dorsum arching was attributable to a velarization gesture, the articulation during the labials must be contrasted to some environment where there is no such gesture. Based on acoustic analyses of [22], stressed Russian [a] has higher first-formant values than any other vowel, consistent with it being the only open vowel in the Russian inventory, and has second-formant values intermediate between the front vowels [i, ɛ] and the back vowels [u, o]. The articulation for [a] was therefore expected be low and devoid of any velar (high) constriction, though a pharyngeal (low, posterior) constriction was possible [7, 18]. The tongue contours of the labials were therefore compared with those of the [a]. If the labials had a velarization gesture associated with them, then the dorsal sections of tongue contours were expected to be higher and more arched than [a].

Figure 2 shows the splines for the contours of the word-final labials and preceding [a] for all three speakers. The tongue contours for [a] are lower than those for word-final labials for all speakers. Increased arching is notable for speakers 2 and 3. Results were the same word-initially for speakers 2 and 3, though speaker 1 had no differences between word-initial labials and the following [a].

However, two aspects of the differences shown in Figure 2 cast doubt on whether these differences were attributable to a velar gesture being associated with the labials but not with [a]. Labial consonants require the lower lip to either make a closure with the upper lip (for [p, m]) or contact the upper teeth (for [f]), while the vowel [a] is open. Jaw height can make a significant contribution to both the closing gesture for the labials and the lowering associated with [a], see, e.g., [5]. Though Figure 2 shows that the tongue was lower for [a] vs. the labial, it is possible that the differences were due predominantly to changes in jaw position. The fact that the shape of the tongue during [a] showed an arching in the posterior part of the contour, though less extreme than the arch found during the labials, is consistent with the differences being attributable to jaw movement.

Despite the acoustics of [a] from [22], it has been claimed that the gestural specification for Russian [a] includes a pharyngeal constriction [18]. It is therefore possible that the dorsal constriction associated with the labials was not part of the specification of the labial segment but was rather a coarticulatory effect due to the adjacent [a]. Another possibility is that the dorsal gesture associated with the labials had a coarticulatory effect on the vowel. In order answer this question, we compared the tongue contours of the [a] in [pam] with the tongue contours of [a] at the beginning of the carrier phrase ([pam] was chosen so that the number of contours for the [a]s would be comparable—the results were the same for [a]s from [fam] and [mat]). The [a] at the beginning of the carrier phrase did not have consonants on either side of it, so the shape of the tongue during that segment should have been devoid of any coarticulatory effects from non-palatalized consonants. While that vowel would normally be reduced to [ɐ] in casual speech [22], speakers 1 and 3 produced the carrier phrase with very deliberate pronunciation of that vowel, and impressionistically it was not reduced. Speaker 2 produced that vowel with some reduction. Figure 3 shows that the arching of the tongue for [a] between the labials (dark blue, comparable to the dark blue in Figure 2) was more retracted than the [a] in isolation, the particular way in which—and degree to which—that retraction was realized varied by speaker.

4. SUMMARY AND DISCUSSION

The present results provide articulatory evidence supporting longstanding claims that non-palatalized consonants in Russian are in fact velarized (and/or uvularized). There was a discernible dorsal gesture

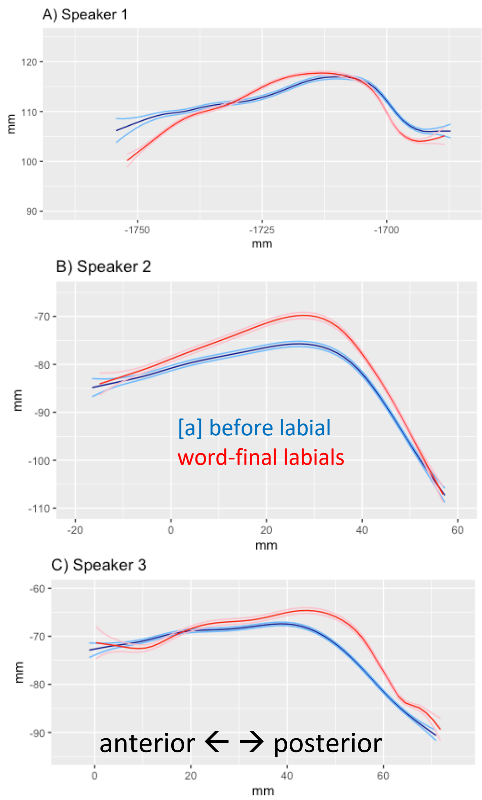


Figure 2: Fitted splines for tongue contours of all word-final labials (red) and adjacent [a]s (blue).

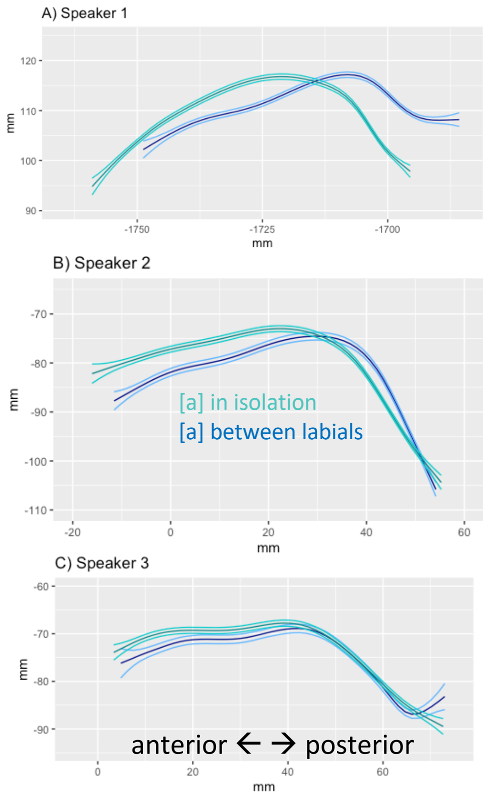


Figure 3: Splines fit to contours of [a] in isolation (cyan) and between labials (blue) from [pam].

associated with three labial consonants [p, f, m], each having a different manner, despite the fact that labials have been described as having less velarization than other segments [17]. There were no regular, significant differences based on the particular segment—and thus manner—across speakers. There were significant differences between word-initial and word-final labials, with the latter having more pronounced (higher) dorsal gestures than the former. These dorsal gestures therefore did not seem to be optional, contra [16]. However, the location and precise shape of the tongue associated with these dorsal gestures did vary by participant, suggesting that the location (e.g., velar vs. uvular) of the constriction may be speaker-specific, rather than segment-specific [18]—at least within labials, regardless of manner.

Comparisons of the tongue shapes for [a] in isolation compared with [a] between labials suggest that the shape of the [a] between the labials was due to coarticulatory effects of the velarization gestures associated with those labials, rather than being part of the gestural specification of [a]. The presence of the more extreme articulations word-finally, while unexpected, is more consistent with the hypothesis that there is a velarization gesture associated with the labial than with that raising being a coarticulatory effect from the adjacent vowel. Carry-over coarticulatory effects on a word-final labial from an [a] would not be expected to be more extreme than the articulations that gave rise to them.

It has long been noted that palatalization of neighboring consonants can trigger vowel allophony in Russian [4, 11, 14]: e.g., the stressed vowel [æ] appears only when both the preceding and following consonants on either side of an underlying /a/ are palatalized: /mʲat/ → [mʲat]/ 'mint' and /matʲ/ → [matʲ] 'mother', but /mʲatʲ/ → [mʲætʲ] 'to knead'. It has also been reported that there is significant coarticulatory influence of palatalization on adjacent vowels, more so in Russian than in English [8]. The present results indicate a strong coarticulatory influence of this secondary velarization gesture on the neighboring vowel as well, comparable to the coarticulation due to palatalization.

It is worth noting that given the idiosyncratic implementation of this gesture across participants, the more holistic quantification of the tongue surface possible with ultrasound was crucially insightful in a way that point-tracking technologies (e.g., electromagnetic articulometry [13]) might not be for this particular phenomenon. The most posterior portion of the tongue, critical for velarization, is difficult to assess with point-tracking technologies. Given the variation in implementation of this dorsal gesture across participants (Figure 2), it would be very difficult to establish in advance what the optimal point of attachment for such a sensor would be.

5. ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from NIH Grant DC-002717 to Haskins Laboratories and the City University of New York.

7. REFERENCES

[1] Avanesov, R. I. 1972. *Russkoe literaturnoe proiznoshenie [Russian literary pronunciation].* Moscow: Prosveshchenie.
[2] Bennet, R., Ní Chiosáin, M., Padgett, J., McGuire, G. 2018. An ultrasound study of Connemara Irish palatalization and velarization. *J. Int'l. Phon. Assoc.* 48, 261–304.
[3] Boersma, P., Weenink, D. 2018. Praat: Doing phonetics by computer. Version: 6.0.21. http://www.praat.org.
[4] Bolla, K. 1981. *A conspectus of Russian speech sounds.* Budapest: Akadémiai Kiado.
[5] Browman, C. P., Goldstein, L. M. 1989. Articulatory gestures as phonological units. *Phonology* 6, 201–251.
[6] Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *J. Acoust. Soc. Am.* 120, 407–415.
[7] Esling, J. H. 2005. There are no back vowels: The Laryngeal Articulator Model. *The Canadian Journal of Linguistics* 50, 13–44.
[8] Evans-Romaine, D. K. 1998. *Palatalization and coarticulation in Russian*. University of Michigan, Ann Arbor.
[9] Fant, G. 1960. *Acoustic Theory of Speech Production.* The Hague: Mouton.
[10] Gu, C. 2014. Smoothing Spline ANOVA Models: R Package gss. *J. Stats. Software* 58, 1–25.
[11] Halle, M. 1971. *The Sound Pattern of Russian.* The Hague/Paris: Mouton.
[12] Heyne, M., Derrick, D. 2015. Using a radial ultrasound probe's virtual origin to compute midsagittal smoothing splines in polar coordinates. *J. Acoust. Soc. Am.* 138, EL509–EL514.
[13] Hoole, P., Zierdt, A. 2010. Five-dimensional articulography. In: Maasen, van Lieshout (eds), *Speech Motor Control: New Developments in Basic and Applied Research.* Oxford: Oxford University Press.
[14] Jones, D., Ward, D. 1969. *The Phonetics of Russian.* Cambridge: Cambridge University Press.
[15] Kochetov, A. 2009. Phonetic variation and gestural specification: Production of Russian consonants. In: Kügler, Féry, van de Vijver (eds), *Variation and Gradience in Phonetics and Phonology.* Berlin: Mouton de Gruyter.
[16] Kochetov, A. 2002. *Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization.* New York, London: Routledge.
[17] Ladefoged, P., Maddieson, I. 1996. *The Sounds of the World's Languages.* Malden, MA: Blackwell Publishing.
[18] Litvin, N. 2014. *An ultrasound investigation of secondary velarization in Russian*. Department of Linguistics, University of Victoria.
[19] Mielke, J. 2015. An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons. *J. Acoust. Soc. Am.* 137, 2858–2869.
[20] Öhman, S. E. G. 1966. Coarticulation in VCV Utterances: Spectrographic Measurements. *J. Acoust. Soc. Am.* 39, 151–168.
[21] Padgett, J. 2001. Contrast dispersion and Russian palatalization. In: Hume, Johnson (eds), *The Role of Speech Perception in Phonology.* San Diego, CA: Academic Press. 187–218.
[22] Padgett, J., Tabain, M. 2005. Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica* 62, 14-54.
[23] R Development Core Team. 2018. R: A language and environment for statistical computing. Version: version 3.2.4. http://www.R-project.org.
[24] Roon, K. D., Dawson, K. M., Tiede, M., Whalen, D. H. 2016. Indexing head movement during speech production using optical markers. *J. Acoust. Soc. Am.* 139, EL167–EL171.
[25] Scobbie, J. M., Wrench, A. A., van der Linden, M. 2008. Head-probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement. In: Sock, Fuchs, Laprie (eds), *Proc. 8th ISSP.* Strasbourg, France: INRIA. 373–376.
[26] Skalozub, L. G. 1963. *Palatogrammy i rentgenogrammy soglasnykh fonem russkogo literaturnogo iazyka [Palatograms and X-ray images of consonant phonemes of literary Russian].* Kiev: Izdatel'stvo kievskogo universiteta.
[27] Stone, M. 2005. A guide to analysing tongue motion from ultrasound images. *Clin. Ling. & Phon.* 19, 455–501.
[28] Tiede, M. K. 2018. GetContours. Version: 1.3. https://github.com/mktiede/GetContours.
[29] Timberlake, A. 2004. *A Reference Grammar of Russian.* Cambridge/New York: Cambridge University Press.
[30] Trubetzkoy, N. 1969. *Principles of Phonology.* Berkeley: University of California Press.
[31] Whalen, D. H., Iskarous, K., Tiede, M. K., Ostry, D. J., Lehnert-LeHouillier, H., Vatikiotis-Bateson, E., Hailey, D. S. 2005. The Haskins Optically Corrected Ultrasound System (HOCUS). *J. Spch. Lang. Hear. Rsrch.* 48, 543–553.