SYLLABLE POSITIONS AS DISTINCT ARTICULATORY PATTERNS: EVIDENCE FROM RUSSIAN*

Alexei Kochetov
Simon Fraser University and Haskins Laboratories

1. Introduction

It has been observed that syllable-initial and syllable-final consonants tend to exhibit different patterns of articulatory organization. These differences, or *syllable position effects*, are often manifested in the relative timing of articulatory gestures and their reduction in magnitude, or the degree of constriction (see Krakow 1999 for a review). Syllable-initial consonants, for instance, tend to show more stable patterns of coordination and "tighter" articulatory constrictions than the same consonants in syllable-final position.

A classic example of these contextual differences involves positional allophones of the American English /l/, namely the syllable-initial "clear" [l] as in 'leap' and the syllable-final "dark" [ɫ] as in 'peal'. Investigations of these allophones using a variety of methods (Giles & Moll 1975, Sproat & Fujimura 1993, Browman & Goldstein 1995, Gick in press) have shown that both involve two articulatory gestures, a tongue tip closure at the alveolar ridge and a tongue body retraction (velarization). The difference between the dark and clear /l/ is in the timing and magnitude of these gestures. First, while the tongue tip gesture is synchronous with the tongue body gesture syllable-initially, the latter follows the former syllable-finally. Second, the relatively consistent achievement of the tongue tip target syllable-initially contrasts with the frequent lack of complete closure syllable-finally, particularly in casual and fast speech.

Articulatory differences in timing and magnitude between syllable-initial and syllable-final allophones are not limited to the lateral. They have been documented for other English consonants: oral and nasal stops (Browman & Goldstein 1995, Keating 1995, Byrd 1996, Turk 1994, Fougeron & Keating 1997), the glides /w/ and /j/, and rhotic /r/ (Gick 1999). The findings suggest that positional variants of English consonants in general reflect distinct patterns of syllable organization at the articulatory level.

* I am grateful to Keren Rice, Louis Goldstein, Elan Dresher, Henry Rogers, Joseph Schallert, and Ron Smyth for help at various stages of this project. The research was supported by a Post-doctoral Fellowship from the Social Sciences and Humanities Research Council of Canada (SSHRC) held at Haskins Laboratories, New Haven, CT, as well as by a Doctoral SSHRC Fellowship held at the University of Toronto. All errors are my own.

Are these findings specific to English only or do they reflect cross-linguistic principles of articulatory organization of the syllable? The current work contributes to the cross-linguistic study of the gestural properties of the syllable by examining positional effects in Russian consonants. The study explores the gestural organization of three consonants, the palatalized voiceless labial stop /p^j/ and the palatal glide /j/, using *EMMA*, the Haskins Electromagnetic Midsagittal Articulometer. The results confirm the hypothesis that the same consonants in these two positions differ with respect to their inter-gestural timing and gestural magnitude.

The paper is organized as follows: In Section 2, I formulate the hypothesis about syllable position effects in Russian and outline specific predictions. The experimental setup, materials, and analysis are discussed in Section 3. The results, presented separately for the inter-gestural timing and for the magnitude of the tongue body gesture, are given in Section 4. This section is followed by the discussion of possible consequences of the findings for perception of phonological contrasts and phonological patterns in general (Section 5).

2. Hypothesis and predictions

As mentioned in the previous section, the study focuses on the positional variation of two Russian consonants, /p^j/ and /j/, as well as of sequences of /p/ and /j/. Labial palatalized stop /p^j/ presents an interesting case for an articulatory study of inter-gestural timing. The consonant is a complex gestural constellation consisting of two coordinated oral gestures, Lips [bilabial, closed] and Tongue Body [palatal, narrow] (Kochetov 2002; see Browman & Goldstein 1989 on parameters of articulatory gestures). An examination of this consonant would allow us to investigate the timing of the two gestures in syllable-initial and syllable-final positions. This timing is compared to the timing of the same gestures in combinations of the consonants /p/ and /j/ (sequences /pj/ and /jp/). In addition, we examine the magnitude properties of the two gestures in the production of syllable-initial and syllable-final consonants /p^j/ and /j/.

The general hypothesis tested in this work is that the Russian consonants in question are different syllable-initially and syllable-finally with respect to inter-gestural timing and their internal gestural properties. That is, the hypothesis predicts that the timing of the gestures lips and tongue body of /p^j/, as well as the magnitude of tongue body in /p^j/ and /j/ differ significantly depending on the position. Based on the findings for English, we may expect the two gestures to be synchronous syllable-initially, and the more closed gesture, the lips, to follow the more open gesture, tongue body, syllable-finally (cf., the timing for the (American) English /l/). In addition, in all positions, the timing of the two gestures of /p^j/ may be different from the timing of the same gestures in sequences of /p/ and /j/, since these are lexically distinct in Russian (Avanesov 1972). Further, we may expect the Tongue Body gesture of Russian /j/ to be reduced syllable-finally (cf., the reduction of the English /j/). The general hypothesis and specific predictions are tested in the current experiment.

3. Methodology

Four speakers of standard Russian, three females and one male (the author), participated in the experiment.[1] Subjects AS and DK were originally from Moscow and subjects NT and AK were from Perm', Russia.[2] The stimuli included nonwords and real Russian words with the consonants /p^j/ and /j/. Nonword utterances, shown in Table 1, contained these consonants syllable-initially (after the vowel /a/ and after the consonants /p/ or /b/ and /t/) and syllable-finally (before the vowel /a/ and before the consonants /p/ and /t/). The utterances combining real words had the same target consonants in the same immediate environments.[3] The stress pattern in both types of utterances was controlled for: both test words of the utterance carried primary stress, for example, *tá[p^j] pápy.* All stimuli were embedded in a carrier phrase: ['ɛtʌ _____ ʌ'p^jat^j] "This is ____ again." Five repetitions of the stimuli were randomized and presented in alternating blocks of nonword and real word utterances in Cyrillic. A total of about 150 articulatory tokens (30 utterances × 5 repetitions) per subject were analyzed.

[1] Some of the data (nonwords with /p/, /p^j/, together with other stimuli for subjects AS, NT, and AK) were used in Kochetov (2002), a study investigating the relation between articulation, perception, and distribution of palatalized stops in various environments.

[2] The differences in accent between the subjects were not considered to be of importance for the study, since the two varieties differ mainly in the quality of unstressed vowels (Ignatkina 1987).

[3] The corresponding real word utterances are combinations (verbs + noun/adjective) of the following Russian words: *shla* (she) "walked," *grab'* "rob" (imp.), *grab* "hornbeam," *brat* "brother," *pjataja* 'the fifth' (fem.), *angela* "angel" (acc. sg.), *pjatogo* "the fifth" (acc. sg.), *padaja* "falling," *pjanyj* "drunk," *bajt* "bite," *vojn* "war" (gen. pl.). The need to control precisely for target consonants and environments took precedence over the semantic plausibility of these word combinations. The word-final labials /b^j/ and /b/ in real word utterances were not expected to differ from the voiceless /p^j/ and /p/ in the nonword stimuli due to their final devoicing (Avanesov 1972).

Table 1. Nonword utterances used in the study

Environment	$/p^j/$	/j/
Onset, V#_V	ta [p^j]apy	ta [j]apy
Coda, V_#V	ta[p^j] apy	ta[j] apy
Onset, C#_V	tap [p^j]apy	tap [j]apy
Coda, V_#C	ta[p^j] papy	ta[j] papy

The data were elicited using EMMA, or the Haskins Electromagnetic Midsagittal Articulometer (Perkell et al. 1992). For each subject, receivers for the articulometer were placed at the following articulators: upper lip (UL), lower lip (LL), lower incisors (as an estimate of jaw movement), and several points on the tongue, namely tongue tip (TT), tongue body 1 (TB1), tongue body 2 (TB2), and tongue dorsum (TD). The movement data were collected at a sampling rate of 500 Hz and the acoustic data at 20 kHz, using the real-time input software Maggie (Tiede et al. 1999). The kinematic data were converted from voltage to distance, calibrated, and corrected for head movement and for any possible shifts of the receivers relative to the transmitters. The analysis of the collected articulatory data was performed using Mavis toolbox (Tiede et al. 1999) for Matlab.

The analysis involved several measurements (see Figure 1). Measurements of inter-gestural timing determined the lag, defined as a period of time between constriction peaks ("max") of the trajectories of the lips and the vertical movement of the tongue body. The lag was considered positive if the tongue body followed the lips and negative if the tongue body preceded the lips. These measurements were made for $/p^j/$ and the sequences /pj/ and /jp/. They were limited to the minimally contrastive nonword and real word utterances with single $/p^j/$ and the sequences of /p/ and /j/. The gestural magnitude measurements determined the constriction peak values ("max" in Figure 1) of the tongue body (the vertical displacement). Tongue body measurements were made for $/p^j/$ and /j/.

The results of these measurements were further tested in a number of separate Analyses of Variance (ANOVAs). Each test involved three between-item factors, Consonant ($/p^j/$ and /j/), Position (syllable onset vs. syllable coda), and Type of stimuli (nonwords vs. real words). The dependent variables included the lag and magnitude of tongue body (vertical displacement). In each case, Tukey HSD post-hoc tests were performed to investigate significant interactions.

In the following section I examine the results of the timing of the lip and tongue body gestures as well as magnitude of tongue body. These are evaluated then based on the predictions made above.

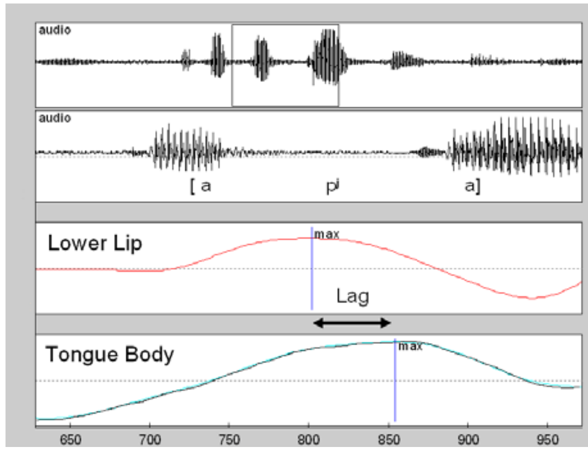


Figure 1. A Mavis display representing the acoustic signal and the trajectories for the lower lip and tongue body during the articulation of /p^j/ in *ta [p^j]apy* (subject NT); the vertical dimension represents the amplitude of the movements; the horizontal dimension represents time (in ms); the labels mark the peaks (max) of the lower lip and tongue body constrictions (vertical displacement); the lag between the constriction peaks of the two gestures is indicated by an arrow.

4. Results

The results are presented in terms of different sets of measurements: inter-gestural timing (Section 4.1) and magnitude measurements for tongue body (Section 4.2). Due to space restrictions, only most relevant significant findings are discussed; the details of the statistical analysis are omitted.

4.1. Inter-gestural timing of Lips and Tongue Body

In this section I determine whether there are positional differences in timing (main effect of Position) between the gestures of lips and tongue body. We also compare the timing of the same gestures in /p^j/ and the two sequences of the consonants /j/ and /p/, that is, /jp/ and /pj/ (Consonant × Position interaction).

The mean values for lag for each subject are presented in Figure 2. Here the values for the lip lag in clusters are given at the bottom and the values for /p^j/ are given at the top. We can see that in the sequences /pj/ and /jp/ there is a substantial lag between the tongue body and lip gestures: for /pj/, the peak of TB follows the peak of the lips by about 80-110 ms (a positive lag); for /jp/, TB precedes the lips by about 30-50 ms (a negative lag). In other words, the two gestures are articulated consecutively, with a minimal overlap, particularly in the /pj/ sequence.

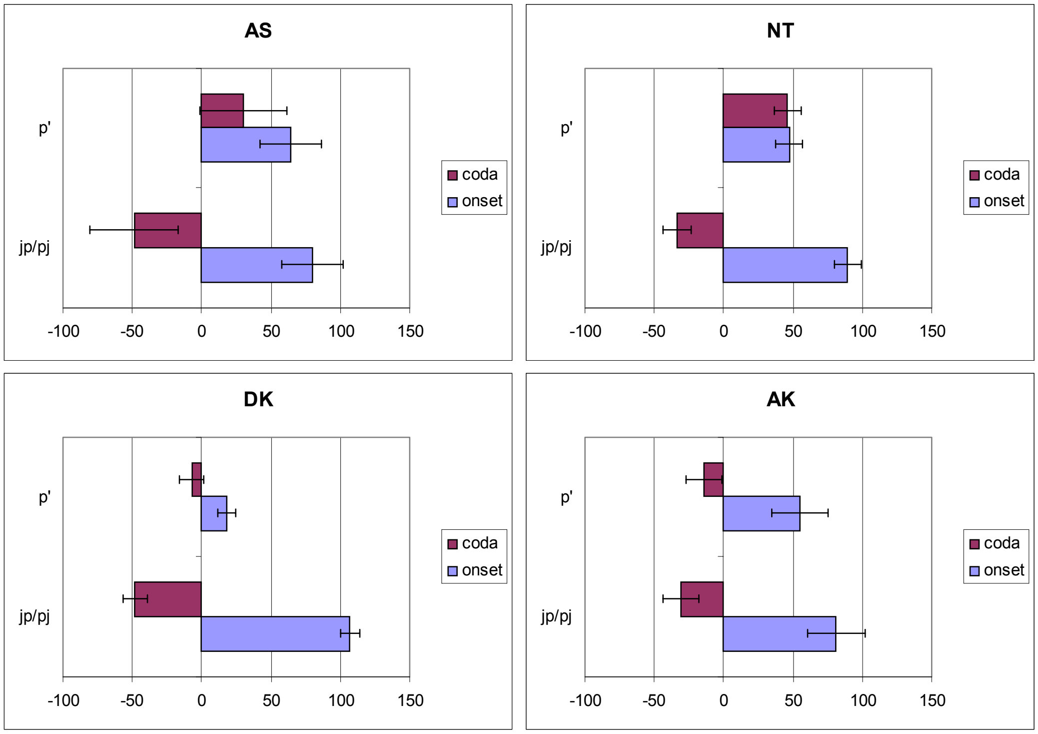


Figure 2. The values of lag between the peaks of lower lip and tongue body during the articulation of /p^j/ (in *ta [p^j]apy* and *ta[p^j] apy*) and the sequences /pj/ (in *ta[p#j]apy*) and /jp/ (in *ta[j#p]apy*) (in ms); four speakers.[4] The lag represents the distance between the peaks of the tongue body and the lips; it is positive when the peak of the tongue body follows the peak of the lips in time and negative when the former precedes the latter.

In this respect the organization of the gestures within the consonant /p^j/ is different from the sequences. The lag between the tongue body and lip gestures is smaller: it is from 20 to 65 ms in onset position and from -15 to 45 ms in coda position. Thus, the two gestures are roughly simultaneous in both positions. However, the timing in the onset position and coda variants of /p^j/ tends to be in the same direction as in the sequences: The lips lag behind the TB for the syllable-initial /p^j/ (for all the four subjects); for the syllable-final /p^j/ the lag is shorter (for subjects AS, DK, and AK); moreover, for two of the subjects the TB slightly lags behind the lips.

These results largely confirm the general hypothesis (see section 2): we do find the syllable effect in the timing of the lips and tongue body of the palatalized labial stop for most of the speakers. The effect is evidenced in the smaller lag in coda position compared to onset position. The overall pattern of the phasing of the

[4] The utterances with /j/ used for subject AK were different and included word-internal /pj/, /jt/, and /jn/ clusters (see Table 1).

gestures of /p^j/ is as follows: Syllable-initially, the achievement of the TB target is somewhat delayed with respect to the achievement of the lips target (it is timed with, or close to, the release of the lips). The two gestures are timed roughly synchronously in syllable-final position (with a possible tendency to earlier attainment of the tongue body).

The observed timing pattern, however, is not the same as expected based on the findings for the English /l/. Recall that the latter showed the synchronous attainments of targets (tongue tip and tongue body) syllable-initially and the lag of the more constricted gesture (tongue tip) syllable-finally. In contrast, the articulation of the Russian /p^j/ involves the two gestures being synchronous in coda position rather than in onset position, and the more open gesture (tongue body) lagging behind the more closed one (the lips).

Compared to the sequences of /j/ and /p/ (/pj/ and /jp/), where the two gestures were phased sequentially, the gestures within /p^j/ were phased relatively simultaneously, with variation with respect to position. Thus, in this respect our predictions were confirmed: these timing patterns are kept distinct in the language. A delay of the tongue body by more than 60-70 ms in onset position may cause a confusion of the palatalized consonant in [p^ja] with a (word-internal) sequence of [p^jja]; similarly, a lead of the TB in coda position by more than 20-30 ms may cause a confusion of the palatalized consonant in [ap^j] with a sequence of [ajp].

4.2. Tongue body magnitude

In this section I examine tongue body magnitude in order to determine whether syllable-initial and syllable-final variants of the consonants differ in these parameters (main effect of Position). We also compare the magnitude of the TB gesture in /p^j/ and /j/ (main effect of Consonant). Recall that the first measurement involved the values of TB raising (vertical displacement) at the trajectory peaks of this gesture for /p^j/ and /j/.

Figure 3 presents the magnitude of the tongue body raising gesture for four subjects. The values for syllable-initial and syllable-final /p^j/ are presented on the left and those for /j/ in the same positions are given on the right.

Note the considerable difference between /p^j/ and /j/ in the degree of TB raising. Both consonants, however, show similar differences in magnitude as a function of syllable position (except for /p^j/ for subject DK). Compared to syllable onset position, /j/ in coda position is reduced by about 1.5 to 4 mm; the reduction of the syllable coda /p^j/ is about 0.5 to 2 mm. Overall, the reduction of the TB of /j/ is more consistent across subjects and is greater in degree than the reduction of the same gesture of /p^j/. As a result, the difference between the two consonants, rather striking in onset position (3 to 6.5 mm), is less apparent in coda position (0.5 to 6 mm).

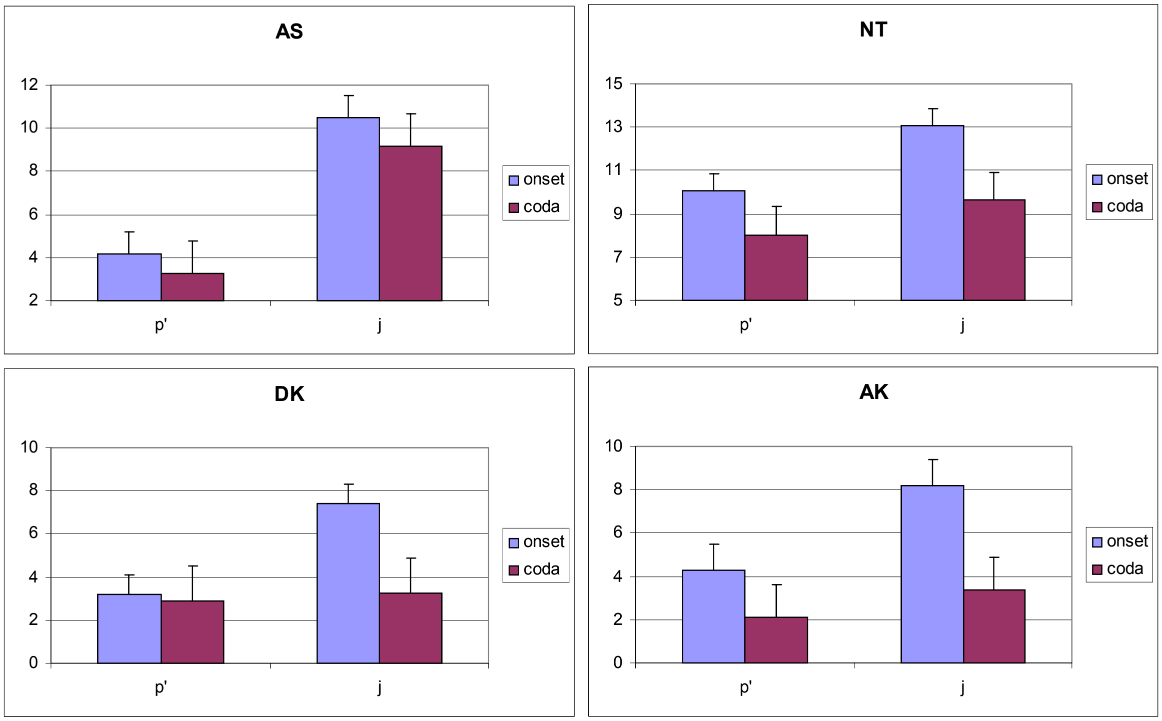


Figure 3. Tongue body raising during the articulation of /p^j/ (labeled as "p'") and /j/ in syllable onset and syllable coda positions (in mm); four speakers. The baseline is arbitrary for all speakers.

The results of the investigation of the tongue body gesture support the general hypothesis: we have identified position-related differences in TB magnitude (TB raising). The results also confirm our specific predictions about the coda reduction of the TB in magnitude: overall, the TB gesture in coda position is considerably reduced in magnitude than the same gesture in onset position. These findings are thus consistent with the results for English /j/, which shows consistent reduction in TB magnitude (Gick 1999).

Another important finding is that the tongue body gestures for /p^j/ and for /j/ were significantly different in magnitude (mainly in the degree of raising) and sometimes in duration. The TB of /j/ was higher than that of /p^j/ and is often characterized by a shorter plateau. The two consonants also differed in the degree of gestural reduction and variation, with /j/ being more affected by the position. This suggests that the TB gesture of /p^j/ is a distinct articulatory structure, with different target (and possibly duration) parameters than those of /j/.

4.3. Summary

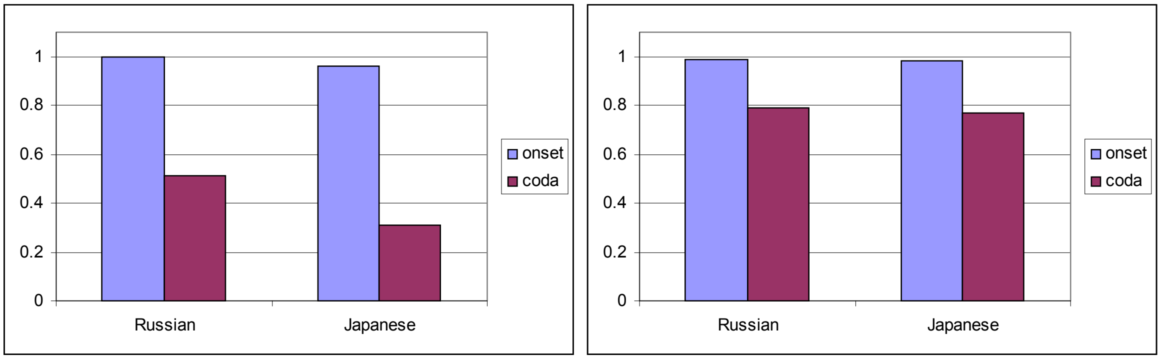
The results of the EMMA experiment support the general hypothesis that syllable position affects the realization of the Russian consonants in question by providing evidence for distinct gestural patterns at least in some syllable-initial and syllable-final positions. The different patterns are manifested in positional differences in timing between the gestures as well as in gestural magnitude. In terms of timing, we found a shorter lag between the lips and tongue body in coda position. In terms of internal gestural properties, we found a reduction of tongue body magnitude in coda position.

5. Consequences for perception and distribution phonological patterns

It is reasonable to expect that the significant differences in timing and magnitude of gestures may influence perception of the consonants in question. Below I discuss how gestural differences specific to syllable position, as well as the degree of overlap of gestures, may affect perception of phonological contrasts and, subsequently, may result in certain characteristic phonological patterns.

A perceptual study (Kochetov 2002) investigated the effects of positional differences in palatalized and non-palatalized stops. The study involved 20 Russian and 10 Japanese listeners who were presented with some nonword tokens with /p/ and /p^j/ from the EMMA experiment described in this paper. The tokens included the two consonants of interest (together with /t/ and /t^j/) in onset and coda positions as produced by subject AK. The task was forced choice identification of consonants (/p/, /p^j/, /t/, /t^j/; the corresponding syllables were used for the Japanese listeners).

Some relevant results are summarized in Figures 4 and 5. As we can see in Figure 4a, perception of /p^j/ as “palatalized” (i.e., identified as /p^j/ or /t^j/ rather than /p/ or /t/) is affected by the position: a perfect, 100% identification of secondary palatal articulation in syllable-initial position by both groups of listeners contrasts with the much poorer identification of this articulation in syllable-final position. Here the identification rate is at chance level for Russian listeners and below chance for Japanese listeners. Figure 4b shows the results of perception of /p^j/ as “labial” (i.e., identified as /p/ or /p^j/ rather than /t/ or /t^j/). The identification of the feature [bilabial] in syllable-final position is still relatively high compared to the perception of [palatal] in the same position (by about 30% for Russian listeners and almost 50% for Japanese listeners). The fact that it is somewhat lower than for the initial consonants reflects the general effect of poorer recoverabilty of consonants in coda position (see, for example, Ohala 1990, Redford & Diehl 1999). As shown in Figure 5, it took longer for listeners of both groups to identify the same consonants in coda position than in onset position.



a. b.

Figure 4. Perception of /p^j/ as (a) palatalized (1.00 = 100% of responses are /p^j/ or /t^j/) and (b) as "labial" (1.00 = 100% of responses are /p^j/ or /p/) in onset and coda positions by 20 Russian and 10 Japanese listeners

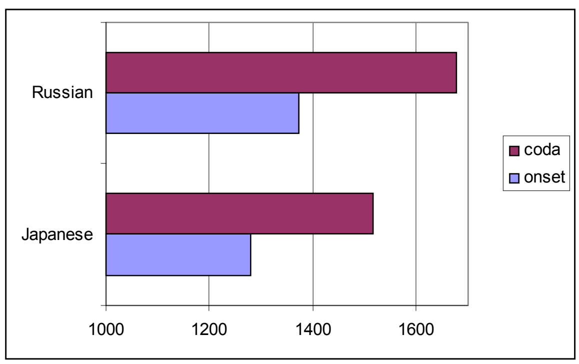


Figure 5. Reaction time during correct identification of /p^j/ in onset and coda positions (in ms) by 20 Russian and 10 Japanese listeners.

In sum, the poorer recoverability of palatalization of /p^j/ (i.e., [palatal]) in syllable coda position can be attributed to the syllable position effects: the characteristic syllable-final coordination of the two gestures of /p^j/ (the near-synchronous achievement of the targets by TB [palatal] and Lips [bilabial]) and the reduction of the TB [palatal] in magnitude. The better recoverability of the lips (i.e., [bilabial]) is likely to be due to its lack of reduction in magnitude (even though it was reduced in duration) together with the more perceptually salient consequences of the lip constriction (see, for example, Surprenant & Goldstein 1998).

Finally, gradient reduction of the tongue body gesture of /p^j/ and its perceptual consequences are likely to be the main diachronic sources of common restrictions against palatalized labials and palatalized consonants in general. As shown in Kochetov (2002), in many languages with contrastive palatalization the distribution of palatalized stops, and particularly, palatalized labials (compared to their non-palatalized counterparts) is limited to syllable onsets. If palatalized stops do occur in the syllable coda position in a language, they tend to be relatively low in frequency and are subject to certain morphological and lexical restrictions.

6. Conclusion

This paper addressed the question of whether and to what extent syllable position effects – articulatory differences between syllable-initial and syllable-final consonants – are language-specific or universal. The hypothesis that Russian consonants would show distinct patterns of timing and reduction in coda position was tested in an articulatory (EMMA) study of Russian labial stop /p^j/ and palatal glide /j/ in syllable onset and coda positions. The results supported the hypothesis: positional allophones of /p^j/ and /j/ do show distinct patterns of gestural organization, which are manifested in different timing and magnitude of the gestures involved. At the same time, certain predictions about relative timing of gestures based on the findings for English were not supported. This suggests that while syllable position differences in general are likely to be universal, their exact realization may be language-specific. It was further shown that the timing and reduction of some gestures have important consequences for the perception and, ultimately, for the distribution of phonological contrasts.

References

Avanesov, R.I. 1972. *Russkoe literaturnoe proiznoshenie*. Moscow: Prosveshchenie.

Browman, C.P. and Goldstein, L. 1989. 'Articulatory gestures as phonological units,' *Phonology* 6, 201-252.

Browman, C.P. and Goldstein, L. 1995. 'Gestural syllable position effects in American English,' In: *Producing speech: Contemporary issues* (for Katherine Safford Harris), eds. F. Bell-Berti and L.J. Raphael. Woodbury, NY: AIP Press, 19-33.

Byrd, D. 1996. 'A phase window framework for articulatory timing,' *Phonology* 13(2), 139-169.

Fougeron, C. and Keating, P. A. 1997. 'Articulatory strengthening at edges of prosodic domains,' *Journal of Acoustic Society of America* 101, 3728-3740.

Gick, B. 1999. The articulatory basis of syllable structure: A study of English glides and liquids. Doctoral dissertation, Yale University, New Haven, CT.

Gick, B. (in press). 'Articulatory correlates of ambisyllabicity in English glides and liquids,' In: *Papers in Laboratory Phonology VI: Constraints on Phonetic Interpretation,* eds. J. Local, R. Ogden and R. Temple. Cambridge: Cambridge University Press.

Giles, S.B. and Moll, K.L. 1975. 'Cinefluorographic study of selected allophones of English /l/,' *Phonetica*, 31, 206-227.

Ignatkina, L.V. 1987. 'Permskie govory,' In:*Interferentsia zvukovykh system,* eds. L.V. Bondarko and L.A. Verbitskaia. St. Petersburg: Izdatel'stvo Leningradskogo Universiteta.

Keating, P. 1995. 'Effects of prosodic positionon /t,d/ tongue/palate contact,' *Proceedings of the 13th International Congress of Phonetic Sciences*, Stockholm, 3, 432-435.

Kochetov A. 2002. *Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization.* New York: Routledge.

Krakow, R.A. 1999. 'Physiological organization of syllables: A review,' *Journal of Phonetics* 27, 23-54.

Ohala, J.J. 1990. 'The phonetics and phonology of aspects of assimilation,' In: *Papers in laboratory phonology I: Between the grammar and physics of speech.* eds. J. Kingston and M.E. Beckman. Cambridge: Cambridge University Press, 258-275.

Perkell, J.S., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I. and Jackson, M. 1992. 'Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements,' *Journal of the Acoustical Society of America* 92, 3078-3096.

Redford, M.A. and Diehl, R. 1999. 'The relative perceptual distinctiveness of initial and final consonants in CVC syllables,' *Journal of Acoustic Society of America* 106(3), 1555-1565.

Sproat, R. and Fujimura, O. 1993. 'Allophonic variation in English /l/ and its implications for phonetic implementation,' *Journal of Phonetics* 21, 291-311.

Surprenant, A. and Goldstein, L. 1998. 'The perception of speech gestures,' *Journal of the Acoustical Society of America* 104, 518-529.

Tiede, M.K., Vatikiotis-Bateson, E., Hoole, P. and Yehia, H. 1999. 'Magnetometer data acquisition and analysis software for speech production research,' *ATR Technical Report TR-H 1999*. ATR Human Information Processing Labs, Japan.

Turk, A. 1994. 'Articulator phonetic cues to syllable affiliation: Gestural characteristic of bilabial stops,' In: *Papers in Laboratory Phonology III: Phonological structure and phonetic form*, ed. P. Keating. Cambridge, England: Cambridge University Press, 107-135.