Transitional vowels and extended sonorants. An acoustic study of Russian coda clusters

Peter Staroverov

petesta@rutgers.edu, Tel: +1 732 9327289
Rutgers, The State University of New Jersey, 18, Seminary Place, New Brunswick, New Jersey, USA

Abstract
Transitional vowels in consonant clusters have played a prominent role in integrating gestural representations within phonology. This article addresses the question of how to diagnose a transitional vowel in a cluster containing a sonorant. It is argued that a vocalic interval may arise in such clusters due to the extension of the sonorant itself.

An acoustic study of Russian word-final stop+liquid clusters is presented. The stop-liquid clusters (as in /tsikl/ 'cycle') were found to contain an interval very similar to a transitional vowel. However the acoustic properties of this interval indicate that it arises via extension and vocalization of the liquid gesture. This finding implies that the extent and pervasiveness of transitional vowel phenomena may need further refinement, especially between a stop and a sonorant.

No vowel interval was found in liquid+stop clusters (as in /polk/ 'regiment'). It is hypothesized that the perceptibility of the liquid would be threatened if a word-final stop-liquid cluster was produced without an audible vocalized transition between the stop and the liquid. On the other hand, the postvocalic liquid is better recoverable. In Russian the realization of the liquid correlates with the sonority composition of the cluster where the liquid occurs.

Keywords
transitional vowels, sonorant extension, consonant clusters, liquids, sonority, Russian

Transitional vowels and extended sonorants. An acoustic study of Russian coda clusters

1. Introduction and background

The realization of consonant clusters sometimes involves a phonetic vowel which behaves differently from full lexical vowels. These *transitional vowels* have received some attention in theoretical as well as in phonetic and modelling literature (Bradley, 2004, 2005, 2007a; b; Davidson, 2006, 2010; Davidson & Stone, 2004; Gafos, 2002; Hall, 2003, 2006; Ramírez, 2002, 2006; Schmeiser, 2009; Shaw & Davidson, 2011). The languages where transitional vowels have been reported include Piro (Matteson & Pike, 1958), Sierra Popoluca (Elson, 1956), Norwegian (Bradley, 2007a). Hall (2003) presents a typological survey of transitional vowels and Davidson (2006, 2010; Davidson & Stone, 2004) analyze production of non-native clusters. However, in some consonant clusters the realization of the cluster itself may lead to the acoustic signal which appears to contain a transitional vowel. These cases need to be carefully distinguished from genuine transitional vowels. This paper contributes one such case and shows how the acoustic properties of the interconsonantal interval may be used to assess the presence/absence of a transitional vowel.

This paper presents an acoustic study of Russian word-final clusters of stops (T) and liquids (L). Although an interval very similar to a transitional vowel was found in the TL# clusters (as in /tsikl/ 'cycle'), the acoustic properties of this interval indicate that it arises via extension and vocalization of the liquid gesture. Thus in cases like Russian a putative transitional vowel is very hard to distinguish from the realization of the cluster itself. Therefore the extent and pervasiveness of transitional vowel phenomena may need further refinement, especially between a stop and a sonorant.

Furthermore, the realization of LT# clusters (as in /polk/ 'regiment') did not show a comparable effect. Thus in the case of Russian sonority composition of a cluster seems to play a role in how this cluster is realized. It is hypothesized that the extension and vocalization of the liquid in word-final rising sonority TL clusters is happening in order to make the liquid more perceptible.

The paper is organized as follows. Section 1.1 reviews the gestural analysis of transitional vowels and section 1.2 gives the background on Russian. Section 2 describes the experiment, the results are reported in section 3 and section 4 discusses the consequences and interpretation of the results. Section 5 concludes the paper.

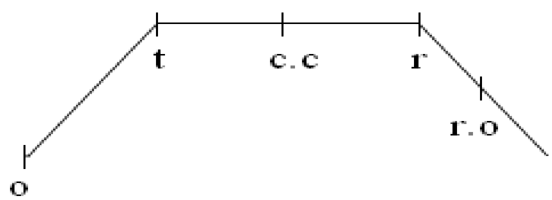
1.1 Realization of consonant clusters and the transitional vowels

The realization of consonant clusters is most commonly represented in terms of gestural phonology (Browman & Goldstein 1986; 1988 et seq.). In this model, each segment is associated with a number of gestures of different articulators. Gestures are characterized by a set of dynamic stages, landmarks. Gafos (2002) proposes that the grammar dictates the coordination relations between gestures of different segments. It is assumed that for each segment there is a head gesture which is the oral gesture (as opposed to velic or laryngeal). Gafos argues that the grammar may require the landmark A of the (head gesture of) segment S_1 to coincide with landmark B of the segment S_2.

(1) shows a schematic representation of a gesture with the gestural landmarks introduced by Gafos (2002). The horizontal line represents the period of time when gestural target is held while the lines to the left and right represent the movement to and away from the target. The gestural

landmarks include ONSET – the point in time when movement towards the target of a gesture begins, TARGET – the point at which a gestural target is achieved, RELEASE – the point when movement away from the target starts, C-CENTER – the point in the middle between target and release, and RELEASE OFFSET – the point when active control of the movement away from the target ceases.

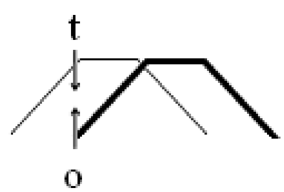
(1) Gestural landmarks: onset (o), target (t), c-center(c.c), release (r) and release offset (r.o)



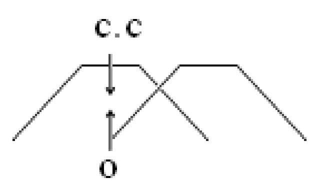
The figures in (2) (adapted from Gafos (2002: 271)) illustrate how two consonant gestures (C1 and C2) can be coordinated.

(2) Examples of coordination relations between consonant gestures

a.



b.

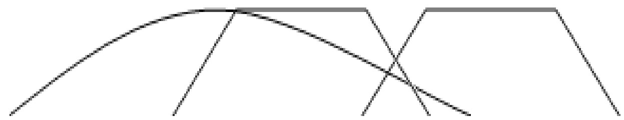


In (2a), the release of the first constricted gesture coincides with the target of the second (the onset of C2 is coordinated with the target of C1). As a result, when C1 is released C2 has already reached its target and no release is audible (Gafos 2002).

In (2b) the two gestures are further apart. The c-center of C1 is coordinated with the onset of C2. In this case there is a time lag between the release of the first consonant and the target of the second consonant. When the two consonants are not homorganic, this time lag corresponds to the period of movement from one target to the other with relatively open vocal tract. An acoustic consequence of this is that a release and a transitional vowel are heard (Davidson 2006; 2010; Gafos 2002). If the two gestures are pulled even further apart, an acoustic release also starts to arise in the homorganic sequences.

Steriade (1990) proposes that many transitional vowels arise when a neighboring vowel gesture overlaps the open transition between the two consonants. Hall (2003; 2006) reports on a typological survey focused on these echo vowels. In these cases the quality of the transitional vowel is similar or identical to the quality of a neighboring vowel. The transitional echo vowels arise more often next to the consonantal gestures which more easily overlap with vowel gestures, such as sonorant consonants. If we represent a vowel gesture with an arc, a transitional vowel arising via gesture sharing can be schematized as in (3) (adapted from Hall (2003)).

(3) Vowel intrusion via overlap with a neighboring V gesture



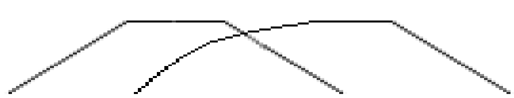
Transitional vowels separated by an obstruent from a neighboring vowel gesture presumably arise as a result of the open transition itself (Gafos 2002). A number of studies report transitional vowels next to obstruents, but the quality of the vowel is rarely systematically studied. The results of the experiment in this paper indicate that a closer look at the vocalic quality may demonstrate that some alleged transitional vowels are better analyzed as consequences of consonant realization.

Schmeyser's (2009) preliminary acoustic study of Guatemalan Spanish tautosyllabic flap + C clusters shows evidence of a transitional vowel which has a schwa-like quality disregard of the quality of the preceeding or the following vowel. The studies of other Spanish varieties (Quilis 1970; 1993; Ramírez 2002; 2006) find that the transitional vowels are echo vowels. Colantoni and Steele (2005) report that the quality of a transitional vowel is significantly affected by both neighboring vowels in Argentinian Spanish and Quebec French obstruent-rhotic clusters. Finally, Davidson (2006, 2010) reports on transitional vowels in English and Catalan speakers' production of word-initial Slavic-like clusters of an obstruent + C. The vowel following the cluster was always the same in Davidson's data, so the amount of V-gesture sharing is not assessed. The transitional vowel within the cluster consistently had a schwa-like quality, although both F1 and F2 were lower than for English schwa.

The present study identifies an interval which is very similar to transitional vowels in Russian TL# clusters. However the acoustic properties and especially formant structure of the alleged vowel indicate that this interval arises via extension and vocalization of the liquid gesture. This gestural pattern is shown in (4) where the vocalization of the second consonant is represented by smoothening the curve. Thus the Russian case illustrates that transitional vowels are often hard to draw apart from details of sonorant realization. A careful analysis of vowel quality is needed in order to postulate a transitional vowel in obstruent-sonorant clusters.

The Russian data also indirectly confirm the claim that gesture sharing is unlikely across an obstruent.

(4) Vowel-like period arising via extension of a sonorant gesture



Russian liquids were found to be extended in coda clusters of rising sonority, but not in coda clusters of falling sonority. /l/ is known to often be realized differently depending on its syllabic position (Sproat & Fujimura, 1993; Gick *et al.*, 2006). However, the distribution of Russian liquid allophones goes beyond the syllable-initial vs. syllable-final dichotomy. Russian liquids are extended in the coda only if they follow another consonant (usually an obstruent). If the liquids occur on their own or precede a consonant, the extension as in (4) is not observed. Furthermore intuitively the intervocalic TL combinations do not exhibit liquid extension. It is hypothesized that the sonority-violating clusters threaten the perceptibility of the liquid (see also Wright (1996, 2004)), which in turn triggers an otherwise unattested phonetic realization.

Gafos (2002: 4.1.2.2) mentions some preliminary evidence that the avoidance of overlap in Moroccan Arabic homorganic clusters is weakened when the two consonants are of different sonority. The evidence in this paper supports the view that the sonority composition of a consonant cluster may affect its gestural realization.

1.2 Russian consonants and clusters

Russian liquids contrast in palatality: /l/ - /lʲ/, /r/ - /rʲ/. The experiment focuses on nonpalatalized liquids in order to better understand the quality of the vowels next to the liquids, although according to native speaker intuitions, palatalized and nonpalatalized liquids behave the same way. In an acoustic study of Russian rhotics Iskarous & Kavitskaya (2010) found that word-final /r/ in Russian can be realized as a full trill (more than one full closure), tap (one full closure) or approximant (one or more incomplete closure). /l/ is less variable in realization and therefore it is used in the experiment. Russian nonpalatalized /l/ is reported be relatively back (presumably due to the palatalization contrast) and to have a dorsal articulatory component (Kochetov, 2005; Proctor, 2009, 2011).

The realization of Russian consonant sequences has been studied across word boundaries and word-initially (Davidson & Roon, 2008; Zsiga, 2000, 2003). However, no studies of Russian consonant clusters in the coda are known. Zsiga (2000, 2003) found that Russian speakers release the first stop in the consonant sequences spanning a word boundary more often than the speakers of English. Davidson & Roon (2008) find that "Russian speakers are fairly consistent in releasing stops before other consonants" (p. 150). Based on these findings, it is expected that Russian word-final clusters will be realized with an audible release.

The realization of word-final clusters is conditioned by their sonority. Only the word-final clusters of rising sonority (5)a intuitively seem to be pronounced with vocalization.

(5) Russian word-final combinations of a stop and a liquid

a. T + L

/bobr/ 'beaver'; /rublʲ/ 'rouble';

/tʲeátr/ 'theater'; /podl/ 'mean.PRED'

b. L + T

/polk/	'regiment'	/kolʲt/	'colt'
/dolg/	'debt, duty'	/skarb/	'stuff, goods and chattels'

Russian is a language with lexical stress and pervasive vowel reduction. In Standard Russian (roughly, the dialect of Moscow and the central segment of the European part of Russia), /a/ and /o/ after non-palatalized consonants are neutralized to [ɐ] in the first pretonic syllable and to [ə] in syllables which are further away from stress (Avanesov, 1972; Barnes, 2006a; b; Bondarko, 1977; Halle, 1959; Jones & Ward, 1969; Lightner, 1965, 1972; Padgett & Tabain, 2005; Scherba, 1912; Timberlake, 2004; Ward, 1975 among others). In (5) and throughout the paper stress is marked in the words which have more than one syllable.

The presence/absence of a lexical unstressed vowel in TL-final words is only marginally contrastive. For example, the second form in the minimal pair in (6) is truly rare.

(6) /pádal/ ‘fall.PST.3SG’ – /padl/ ‘infml. bad man.GEN.PL’

In general, the words which have a word-final TL are fairly infrequent in Russian compared to words with final clusters of falling sonority, both in token frequency and in type frequency of words which end in the relevant cluster in the citation form (Proctor, 2009). Many of the words whose citation forms end in TL are borrowings, but a lot of them have been fully assimilated. There are also native words which end in TL in the citation forms. These frequency facts led Proctor (2009) to conclude that words with sonority-violating final clusters "can be seen as historical anomalies which do not reflect syllable structure preferences in Modern Russian" (p. 127).

However, even though word-final TL clusters are statistically underrepresented, the grammatical knowledge of Russian has to include the knowledge of how the word-final TL clusters are pronounced. The combined token frequency of all tokens with word-final Stop-Liquid sequences is 290.13 items per million (based only on tokens with frequency 1 item per million or more in Sharoff (2002)). Thus these words are heard with some frequency in Russian speech. Furthermore, there are grammatical patterns which yield these kinds of sequences. These patterns include the formation of genitive plural and predicative forms of adjectives. Russian speakers have to know how to pronounce word-final TL sequences and this knowledge will be addressed in the experiment.

To sum up, word-final TL sequences in Russian appear to be realized with a phonetic vowel-like interval whereas LT sequences are not. The same holds for word-medial coda sequences although there is only a handful of words with word-medial TL codas, most notably those with the prefix /kontr-/ 'counter-' (e.g. /kontrrazv^{j}édka/ 'counter-intelligence', /kontrr^{j}evol^{j}utsiónnij/ 'counter-revolutionary') and the derivatives of the root /bodr/ 'cheerful' (e.g. /bodrstvovat^{j}/ 'be awake').

This article reports on an experiment which confirms the intuitions about the realization of word-final TL and LT and addresses the status of the vocalic interval heard in the TL sequences.

2. Method

This section describes the nonce-word production task which was used to assess the experimental questions.

2.1 Stimuli

The acoustic study was designed to address the following questions.

1. Is there a vocalic interval in the word-final TL clusters?
2. Is there a vocalic interval in the word-final LT clusters?
3. If there is a vocalic interval in TL clusters, what is the most appropriate interpretation of its status? Is it an epenthetic vowel, a transitional vowel, or a part of consonant realization? Could it arise from gesture sharing with the preceding vowel?

In order to address the first two questions, the experiment compares four kinds of underlying word-final sequences in a nonce-word production task: TL#, TVL#, LT#, LVT#. The comparison between TL and TVL on the one hand and LT and LVT on the other hand addresses whether there is a vocalic interval in TL and in LT.

The comparison between TL and TVL condition addresses the status of the vowel in TL (if present). If the vowel in TL is realized the same as a lexical vowel, the vocalic interval is epenthetic. On the other hand, if there are acoustic differences between the TL and TVL conditions, the vocalic interval in TL may be better interpreted as a transitional vocoid.

Finally, the quality of the vowel in front of the stop was varied in order to address the gesture sharing hypothesis. If the vocalic interval arising in the TL condition is a result of gesture sharing with the preceding vowel, we expect that its quality will exhibit a stronger correlation with the quality of the preceding vowel than that of a lexical schwa.

The stimuli were nonce-words constructed to look like possible Russian words. The use of nonce-words allows to control for the context in which the word-final cluster appears.

The final sequence of the stimuli orthographically was one of the following: "kl", "kal", "lk", "lok" (all stimuli were in Russian orthography). All the stimuli contained two syllables in front of the target sequence with stress marked on the second syllable of the word (the conventional Russian stress marker familiar to all the speakers was used). It should be noted that Russian post-tonic vowel reduction leads to complete neutralization of /o/ and /a/ (Barnes, 2006a; b; Padgett & Tabain, 2005); thus the vowel in /-kal/ words was expected to be the same as in /-lok/ words. /-kal/ and /-lok/ were used to increase the word-likeliness of the stimuli (based on corpus frequency from Sharoff (2002)).

The quality of the stressed vowel was varied between /i/ and /o/ since these vowels differ in both F1 and F2. Furthermore, the words with stressed /i/ contained a palatalized consonant in front of /i/ while the words with stressed /o/ contained a non-palatalized consonant. This way the backness difference between the two vowels was maximized since /i/ is known to have a more centralized realization after non-palatalized consonants (Padgett (2001) and references therein). This difference was introduced in order to assess the gesture sharing hypothesis. If the vowel interval in the TL condition arises via gesture sharing with the preceding vowel, we would expect a stronger correlation in quality of the vowels in the TL condition than in the TVL condition.

The first syllable of all stimuli was manipulated in order to avoid turning the TL vs. TVL and LT vs. LVT stimuli into minimal pairs. The first syllable was one of /na-/, /za-/ and /po-/ which are common Russian prefixes.

Table 1 summarizes the experimental conditions and gives sample stimuli for each condition (see Appendix A for a full list of stimuli). /i/ was preceded by palatalized consonants, which is indicated by a palatality symbol in the table.

/l/ was chosen as the target liquid because the realization of the other Russian liquid /r/ varies greatly across tokens. Iskarous & Kavitskaya (2010) found that word-final /r/ in Russian can be realized as a full trill (more than one full closure), tap (one full closure) or approximant (one or more incomplete closure). A pilot study confirms great variability in realization of Russian word-final /r/. Given this variability, it would be hard to consistently detect and analyze the vocalic interval in word-final Tr and rT sequences since the different realizations of /r/ may affect the acoustics of preceding vowel differently. /k/ was chosen as the stop in order to maximize the word-likeliness of the stimuli based on the frequency information in Sharoff (2002).

Table 1 Experimental conditions and sample stimuli

Type of sequence	Preceding V	Final sequence	Transliterated example	Transcribed example
-TL	o	-okl	zadókl	[zɐdókᵊl]
	(ʲ)i	-ʲikl	nafʲíkl	[nɐfʲíkᵊl]
-TVL	o	-okəl	zagókal	[zɐgókəl]
	(ʲ)i	-ʲikəl	pozʲíkal	[pɐzʲíkəl]
-LT	o	-olk	nazólk	[nɐzólk]
	(ʲ)i	-ʲilk	posʲílk	[pɐsʲílk]
-LVT	o	-olək	nanólok	[nɐnólək]
	(ʲ)i	-ʲilək	pokʲílok	[pɐkʲílək]

There were 5 items in each condition for each preceding vowel. Each stimulus was repeated 4 times for a total of 160 tokens per speaker (4 conditions x 2 preceding vowels x 5 items x 4 repetitions). The stimuli were interspersed with 40 fillers, each repeated 4 times. All the fillers were orthographically disyllabic or trisyllabic with stress on the second syllable. The fillers ended in the orthographic sequences: "okla", "ikla", "olka", "ilka", "ol", "il", "ok", "ik". The stimuli and fillers are given in Appendix A.

2.2 Participants

The participants were ten native speakers of Russian aged between 20 and 50. Six of the participants were male and four were female. All of the participants were born and raised in Moscow area except one who was born in Dagestan (Southern Russia) but spent a substantial amount of time living and working in Moscow. One other participant reported occasional stuttering, but he never stuttered during the experiment. The participants received no compensation.

2.3 Procedure

The stimuli and fillers were placed in a frame sentence (7) and presented to the speakers in Russian orthography with stress marked on the target word[1]. Within this sentence the target word received focus since it contained new information.

(7)	nɐjdʲítʲɪ	slóvə	X	f	sləvɐrʲé
	find.IMP.PL	word.ACC	X	in	dictionary.LOC
	'Find the word X in the dictionary'				

The experiment contained a short training session (2 trials) needed to ensure that the speakers understood the instructions. After that, the set of sentences was presented to the speakers 4 times on a computer screen, each time in a different random order (the stimuli were randomized using a Microsoft Word macro in Appendix B). The participants were instructed to read the sentences out loud at a normal and preferably constant speech rate. There was a short break after the first two

[1] There are no alternations where a word-final TL cluster alternated with a TVL sequence and thus it was assumed that the speakers would interpret the TL-final stimuli as having an underlying cluster. In general, it was assumed that the speakers postulate the underlying form faithful to the orthographic input.

repetitions. Whenever the speakers mispronounced the target word, they were prompted to repeat the whole sentence again.

Post-experiment feedback was elicited in order to find out whether the speakers could guess the goal of the experiment. 4 out of 10 participants showed some awareness that the experiment was comparing the vowel in TL and TVL conditions.

The recordings were performed in a quiet room in Moscow onto a laptop computer using a Shure SMC-10 directed head-worn microphone and an Edirol UA-25EX preamplifier. The speech was recorded as mono sound and digitized as WAV files with a 32-bit sample bit depth and a sampling rate of 44.1kHz using Audacity.

2.4 Annotation

Within each stimulus, the target interval was manually annotated in Praat (Boersma & Weenink 1999-2011), as exemplified in Figure 1. The annotation was based on the waveform and the wideband spectrogram. The spectrograms were rendered with the default setting of window length 5 ms, dynamic range 50dB and time step 1.25 ms.

In the TL and TVL conditions, the target interval was the period from the offset of the stop burst to the end of the periodic sound of /l/ (the target word was followed by a voiceless fricative within the carrier phrase). For words ending in /-lk/ and /-lok/ the target interval included the stressed vowel and lasted until the beginning of the stop closure (i.e. it was Vl in the VLT condition and VlV in the VLVT condition). The target interval was defined this way because it was not possible to draw an objective boundary between /l/ and the neighboring vowels (Turk *et al.*, 2006).

The stressed vowel was annotated in the TL and TVL conditions and the beginning of the stressed vowel was annotated in LT and LVT as part of the target interval annotation. The beginning of the stressed vowel was assumed to coincide with the beginning of periodic sound after voiceless stops and fricatives. After voiced stops and fricatives, the beginning of the stressed vowel was annotated at the onset of formant structure. Finally, after nasals and /r/, the beginning of the stressed vowel was assumed to coincide with the rise of energy in higher formants. The offset of the stressed vowel in TL and TVL conditions was annotated at the end of periodic waves with clear formant structure for /k/ closure.

Finally, the vowel+glide sequence [ɐj] of the first syllable within the carrier phrase was annotated for all tokens from the rise of energy in higher formants to the end of formant structure or to the point of abrupt drop of energy in higher formants. The first syllable of the phrase served as a baseline for intensity measurements.

Figure 1 shows the annotation of a token of /pok^{j}ikl/ by speaker MT. The vowel in parentheses in the transcription corresponds to the hypothesized vocoid in /kl/.

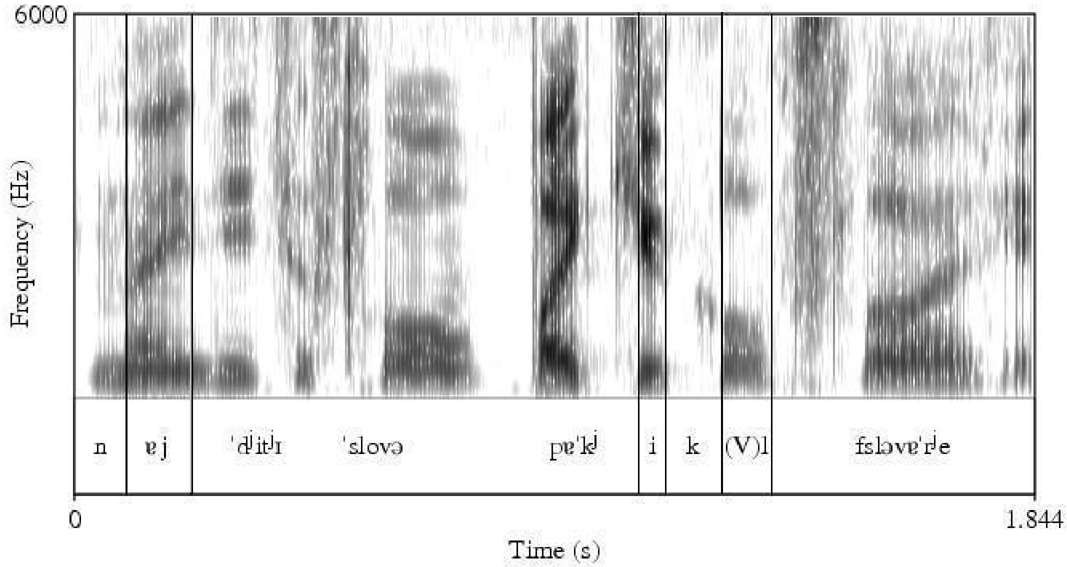


Figure 1. Annotated spectrogram of the first token of /pokʲikl/, speaker MT.

2.5 Acoustic analysis and statistics

All acoustic measurements were taken using Praat (Boersma & Weenink 1999-2011). The duration of the target interval was taken in all four conditions. One token was excluded from all measurements since it contained a mispronunciation and the speaker did not repeat the sentence properly.

In order to assess the energy distribution within the target interval, the interval was broken down into 20 windows and mean intensity (dB) within each window was measured with the minimum periodicity frequency set at 100Hz. The intensity was then normalized by subtracting the mean intensity (dB) of the baseline interval. For a number of tokens, it was not possible to identify the boundaries of the baseline interval ([ɐj] in the beginning of the phrase). Within the carrier phrase the baseline interval was followed by a /d/ which was ocasionally realized as an approximant instead of a full stop. When it was impossible to identify the boundaries of the baseline interval, the relevant tokens were excluded from intensity measurements. Overall, only 13 tokens out of 1600 (0.8%) were excluded for this reason.

The formants were measured at the midpoint of the stressed vowel and at the end of the third window within the target interval in the TL and TVL condition. Since the target interval included both the liquid and the vowel, the end of window 3 (i. e. the 3/20 of the interval duration) was chosen as the point representative of the vowel since the considertaion of the averaged formant contours revealed that at this point both F1 and F2 in both conditions reached a steady state. The formant measurements were taken using linear interpolation Burg LPC with a time step of 10ms, window length of 25ms, and pre-emphasis of 50Hz. The maximum formant was set to 5500 Hz for female speakers and to 5000 Hz for male speakers.

The target word was focused within its sentence. As a result, the speakers ocasionally produced a strong burst in front of the stressed vowel or were pronouncing the stressed vowel with particularly high amount of effort and high airstream. Since a head-worn microphone was used, some of these productions were recorded with noise on the stressed vowel and the corresponding tokens were discarded from formant measurements. The stressed vowel formants were only taken in TL and TVL conditions (see section 3.4) and the total number of tokens excluded was 16 - 2% of 800 tokens in those two conditions.

In addition, in 4 of the TL tokens (out of the overall 400 tokens in that condition) the target (V)L sequence only appeared as a voice bar – these tokens were also excluded from formant measurements.

The statistical analysis was done in R (R Development Core Team, 2011). Linear mixed model (Bates, 2005; Baayen *et al.*, 2008; Baayen, 2008) was run using `lme4` package (Bates *et al.*, 2008). The `lme4` package does not automatically compute p-values because the exact procedure to calculate degrees of freedom has not been discovered. Therefore, the significance of the fixed effects was checked against the 95% confidence intervals calculated by the Markov chain Monte Carlo method using the `pval.fnc()` function of the `languageR` package (Baayen, 2009). Since the number of tokens in each condition was reasonably large the parametric statistics were used.

3. Results

3.1 Duration

The duration of the target interval in TL condition was compared to the duration in TVL condition using a linear mixed model with condition as fixed factor and speaker as a random factor. An analogous comparison was made for LT and LVT conditions.

Figure 2 compares the target interval in the LT and LVT condition based on the first tokens of /natʲilk/ and /pokʲilok/ by speaker MT. In the LT and LVT conditions the target interval embraced the period from the onset of the stressed vowel to the beginning of the stop closure (i.e. the target interval was [il] in the production of /natʲilk/ and [ilə] in the production of /pokʲilok/).

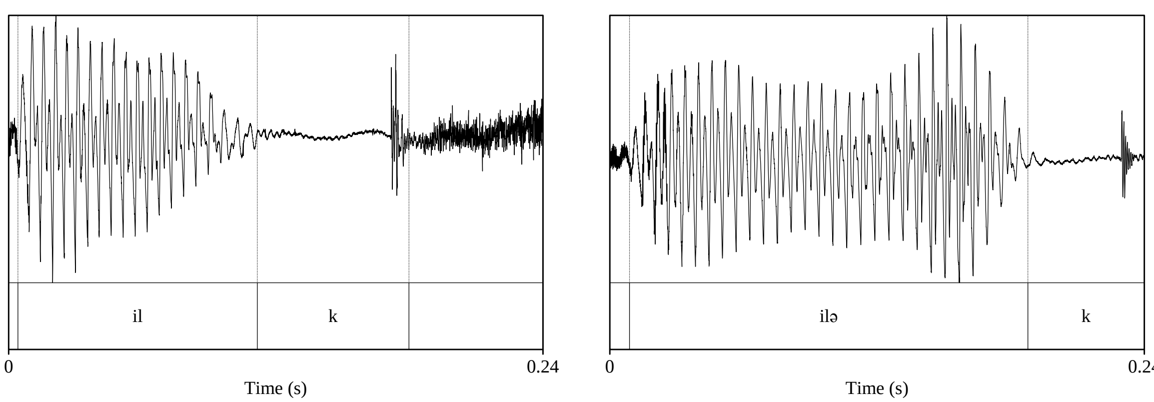


Figure 2. Target interval plus the following stop from speaker MT's first productions of /natʲilk/ (left panel) and /pokʲilok/ (right panel)

On average the target interval in the LVT condition was significantly longer than the target interval in the LT condition. Here and throughout the paper the mean is reported together with a margin of error for a 95% confidence interval in parentheses: 202.5 ms. (±2.9ms) vs. 141.6 ms. (±2.4ms); $p<0.001$.

Figure 3 shows the target interval in the TL and TVL condition based on the third tokens of /nadʲikal/ and /nafʲikl/ by speaker MT. In the TL and TVL condition the target interval corresponded to the period from the beginning of periodic sound after the stop burst up to the end of periodic sound (recall that the target word was followed by a voiceless fricative within the carrier phrase).

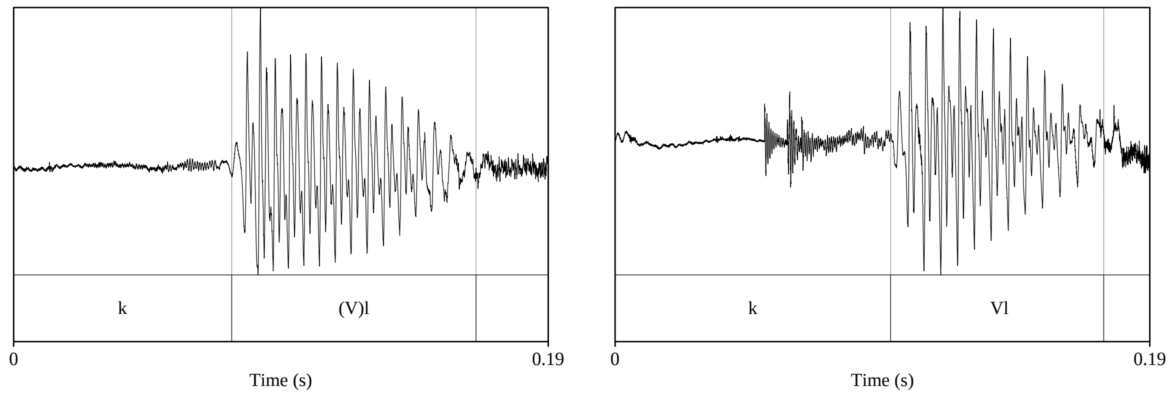


Figure 3. Target interval plus the preceding stop from speaker MT's third productions of /nadʲikal/ (right panel) and /nafʲikl/ (left panel).

On average the target interval in TVL condition was longer than the target interval in TL condition. However the absolute difference was about 7ms., close to the just noticeable difference (Fujisaki *et al.*, 1973; Klatt, 1976): 108.4 ms. (±2.8ms) vs. 100.8 ms. (±3ms); p<0.001.

To examine the small duration difference between TL and TVL more closely, the individual speaker patterns were considered. Table 2 shows the duration differences between TVL and TL conditions by speaker. The significance of the differences was assessed using a two-sided non-paired t-test. A negative difference indicates that TL interval was longer on average. The significance values in Table 2 reflect Bonferroni adjustment.

Table 2 Duration differences between the target interval in TVL and TL condition by speaker. Awareness of the TL vs. TVL contrast is shown based on post-experiment feedback.

Speaker	Duration diff.: TL - TVL (ms)	t.test (Bonf. adjusted)	Aware of the contrast
EK	-8.6	t=0.93; n.s.; df=77	
ES	-2	t=0.55; n.s.; df=61	
OE	-0.3	t=0.07; n.s.; df=74	
KG	2	t=-0.47; n.s.; df=78	
DS	2.5	t=-0.6; n.s.; df=74	
BN	9.7	t=-3.2; p<0.005; df=74	+
PC	12.8	t=-3.4; p<0.005; df=78	
MT	16.7	t=-5.1; p<0.001; df=75	+
KR	19.9	t=-3; p<0.005; df=68	+
EP	23.4	t=-4; p<0.001; df=77	+

Only five speakers out of ten showed a duration difference between the two conditions which was significant on a t-test. TL condition had a longer vowel for all these speakers. Four of these five speakers showed some awareness of the contrast between TL-words and TVL-words on post-experiment feedback. Out of the remaining 5 speakers 4 showed almost no absolute difference in duration between conditions (for these speakers the TVL-TL difference was between +2.5 and -1 milliseconds) and one speaker showed a reversal – the vowel interval in the TL condition was longer by 8.6 milliseconds (n.s.).

The duration difference between TL and TVL condition could be a task effect since the difference was most apparent for the speakers which were aware of the contrast. Overall, the duration patterns are consistent with the hypothesis that there is no vocalic interval in the LT condition, while in the TL condition there is a vocalic interval.

3.2 Energy distribution and presence of a vowel

In order to assess the presence of a vocalic interval in the TL and LT conditions, the distribution of energy within the target interval was examined. Figure 4 shows the schematized distribution of energy in the LT and LVT conditions. The intensity measurements were normalized by subtracting the mean intensity (dB) of the initial syllable [ɐj] sequence. Thus 0 on these graphs corresponds to the interval being as loud as the first vowel of the carrier phrase. As stated in section 3.1, the graph for the LVT condition corresponds to a longer period than the graph for the LT condition.

The energy distribution in the LT and LVT conditions is different. The intensity graph for the LVT condition has a slight rise in intensity after the first peak whereas the graph for the LT condition shows a gradual fall after the first peak. The first peak corresponds to the stressed vowel and the rise after that in LVT condition to the realization of the underlying vowel. The fact that there is no such rise in the LT condition shows that there is no vocalic interval there between the liquid and the stop.

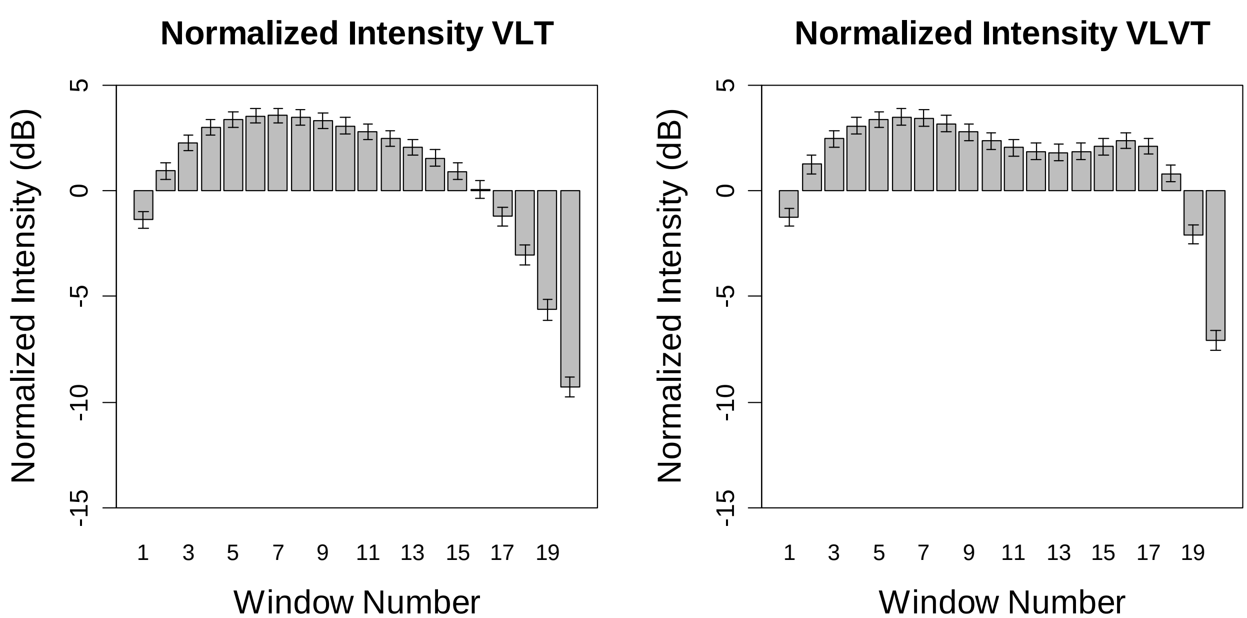


Figure 4. Normalized mean intensity distribution in LT and LVT conditions

Figure 5 shows the energy distribution in the TL and TVL conditions. In this case intensity contours are similar, and there is only one intensity peak suggesting that a vocalic interval is present not only in TVL condition but also in TL. Even though the target interval in the TL condition generally seems to be quieter than in the TVL condition, there is still a period which is as loud as the vowel in the first syllable of the carrier phrase (see section 3.3 on intensity differences).

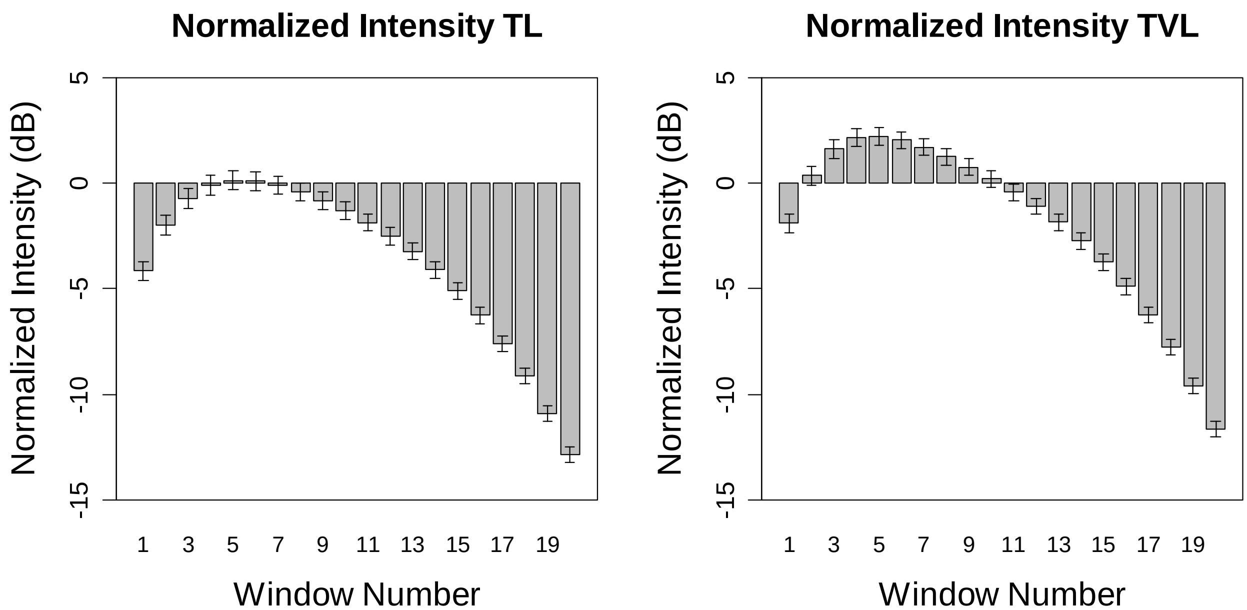


Figure 5. Normalized mean intensity distribution in TL and TVL conditions

To sum up, the acoustic data suggest that there is a vocalic interval in the TL condition, but not in the LT condition thus confirming that the realization of the cluster depends on its sonority composition.

3.3 Acoustics of the vowel in TL

In order to find out whether the vowel in TL condition was acoustically different from the vowel in TVL condition, duration, intensity, and formant structure next to the beginning of the two intervals was compared.

As described in section 3.1, the target interval in the TL condition was found to be shorter, but only by a small barely perceptible amount of time. The small difference in duration may be attributed to a task effect since 4 out of 5 speakers who produced a significant difference, reported awareness of TL vs. TVL contrast (Table 2).

Mean intensity within the 4th window of the target interval was compared between the two conditions using a linear mixed model with condition as a fixed factor and speaker as a random factor. The 4th window was chosen since this was the earliest time when the energy distribution reached a steady state (Figure 5). It was found that intensity was lower in the TL condition: 56.3 dB (±0.52dB) vs. 58.7 dB (±0.49dB) on average ($p<0.001$).

Table 3 summarizes the formant differences between the two conditions. The vowel in TL condition had both a lower F1 and a lower F2 than the vowel in TVL. On the other hand, F3 was slightly higher in the TL condition than in the TVL condition.

Formant measurements at the beginning of window 4 of the target interval were compared using a linear mixed model. Condition (TL vs. TVL) and the value of the relevant formant (that is F1 for F1, F2 for F2 etc) at the midpoint of the stressed vowel were fixed factors and speaker was a random factor. The p-value estimates in Table 3 were obtained using the Markov Chain Monte Carlo method (see section 2.5).

Table 3. Formant differences between TL and TVL condition. Negative difference means that the relevant formant is lower in the TL condition.

	F1 (Hz)	F2 (Hz)	F3 (Hz)
TL condition	423 (±5)	896 (±14)	2751 (±35)
TVL condition	486 (±8)	1006 (±14)	2670 (±35)
Difference and p-value	**-63, p<0.001**	**-110, p<0.001**	**81, p < 0.001**

In addition to the effect of condition, F1 and F2 of the target interval were significantly affected by the respective formants of the stressed vowel. These results will be discussed in section 3.4.

In general, the formant values for each speaker followed the general pattern in Table 3. The absolute differences were either relatively minor or in the expected direction. The F1 and F2 difference was significant on a t.test for all speakers except for two males: KR and ES. The F3 difference was significant on a t.test for four speakers, marginal for one speaker (KR, p=0.009)[2] and insignificant for five speakers: BN (male), KG (male), OE (female), MT (female), EK (female).

Overall, the vocalic interval in the TL condition is acoustically different from the vowel produced in the TVL condition. The vowel in TL is quieter and possibly slightly shorter, and has lower F1 and F2 but higher F3 than the vowel in TVL. These acoustic differences do not support the hypothesis that the vocalic interval in TL is a result of phonological epenthesis. If a vowel segment was epenthesized in TL, we would expect it to be reduced to the same quality as the underlying /a/ and /o/. The most influence on the quality of the vocalic interval is exerted by /l/ as witnessed in F2 and F3. See section 4.1 for further discussion of the formant differences.

3.4 Similarity to the preceding vowel and gestural status of the vowel in TL

If the vocalic interval in TL arises through gesture sharing with the preceding vowel (Hall 2003), we expect that the quality of the vowel in TL will be substantially affected by the quality of the vowel in the preceding syllable. It is also expected that the quality of the vowel in the TL condition should correlate more with the quality of the preceding vowel than in the TVL condition. These predictions are tested with our data since the quality of the stressed vowels varies between /i/ and /o/. The gesture sharing hypothesis also predicts that the vowel in the TL condition will to some extent reflect individual variation in the pronunciation of each stressed vowel.

The linear mixed model analysis in section 3.3 revealed that in both conditions F1 and F2 of the target vocalic interval were significantly affected by the respective formant values of the stressed vowel. In order to assess whether the stressed vowel affected the target vowel differently in different conditions, the interaction between condition and stressed vowel formants was added as a factor to the linear mixed model. The interaction of condition and formants of the stressed vowel was only found to significantly affect F1 of the target vocalic interval. In addition, in the new model condition was no longer significant for F1.

The gesture sharing hypothesis predicts that the interaction between the stressed vowel and the target vocalic interval will be stronger in the TL condition. Post-hoc correlation tests were performed to address the difference in the interaction across conditions. The speakers have

[2] Since multiple pairwise comparisons are made the significance value has to be adjusted and 0.009 should not be considered significant.

different vocal tract lengths resulting in different formant ranges and hence the raw formant values may be correlated just because they come from the same speaker. In order to address this confound, in the post-hoc tests the formant values were normalized by subtracting the mean formant value for the same interval within each speaker.

Figure 6 illustrates the overall correlation between the quality of the vowel after the stop and the quality of the stressed vowel in the TL and TVL conditions. This figure includes the data on tokens with both stressed /i/ and stressed /o/. Normalized formant values are reported; hence 0 corresponds to the mean formant value within each speaker.

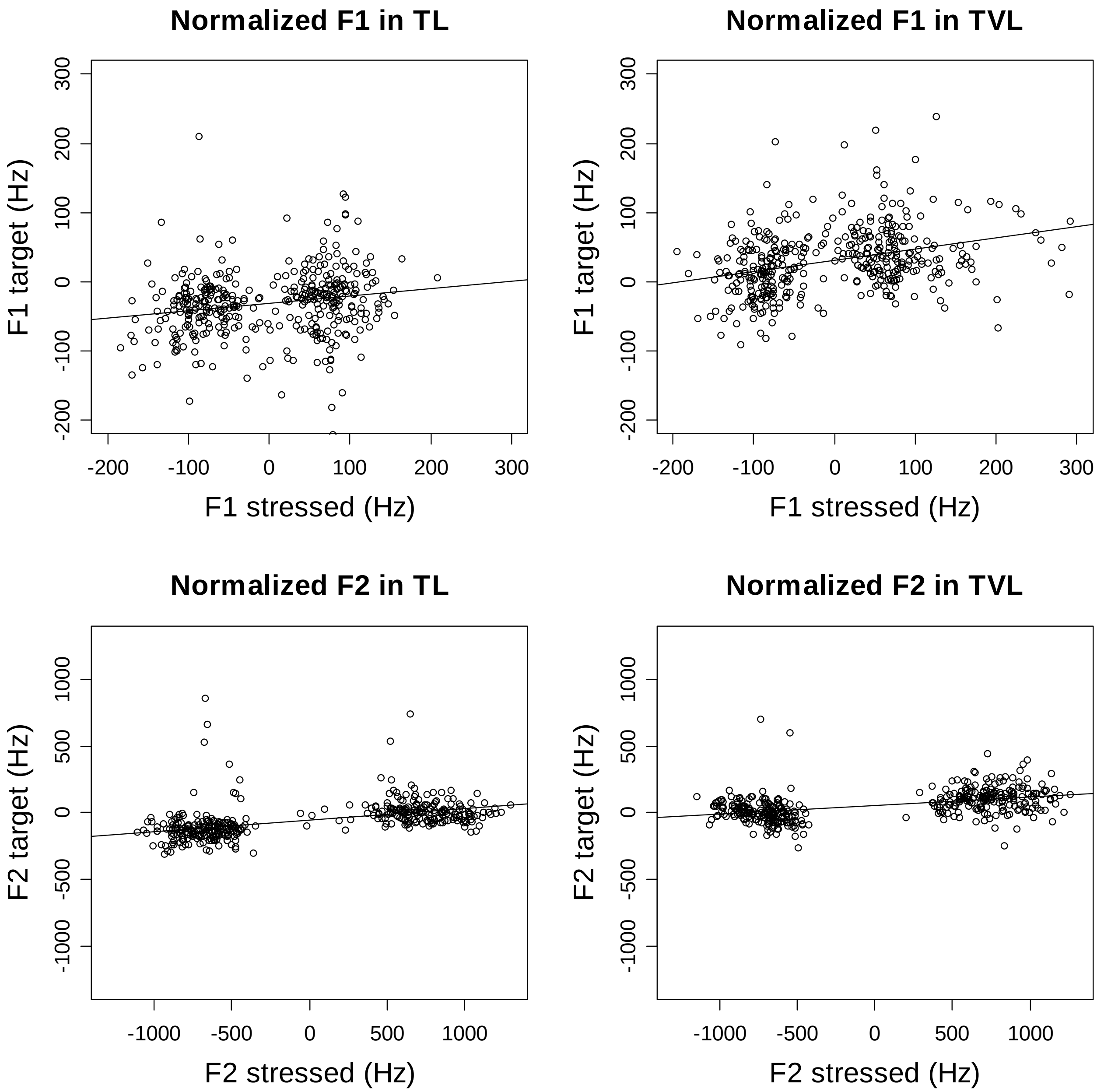


Figure 6. Normalized formant correlation scatterplots in the TL and TVL conditions.

Table 4 shows the Pearson product-moment correlation for F1 and F2 in the two conditions[3].

Table 4. Pearson product-moment correlations for formant values of the stressed vowel and the target vowel depending on condition.

	F1 in TL condition	F1 in TVL condition	F2 in TL condition	F2 in TVL condition
r	0.19 (p<0.001)	0.33 (p<0.001)	0.47 (p<0.001)	0.46 (p<0.001)

[3] In a study of vowel quality in Russian VCV sequences Purcell (1979) only finds evidence of V_1-V_2 coarticulation for F1. However, the current experiment revealed a significant correlation for both F1 and F2 (Table 4). The current study is different in that the lexical vowel in question is a reduced schwa rather than a full vowel. A reduced vowel may be more susceptible to coarticulation.

If the vocalic interval in the TL condition arises through gesture sharing we would expect that there would be a stronger correlation between its quality and the quality of the preceding vowel than in the TVL condition. However, Figure 6 and the correlation figures in Table 4 indicate that the correlation is actually weaker in the TL condition for F1. Using Fisher's z-transformantion the z-score for the correlation difference for F1 is -2.08, $p<0.05$. For F2 the correlation is not significantly different across conditions (z-score is 0.22, n.s.).

Table 5 lists the correlation scores counted separately for tokens with each stressed vowel within each condition. These correlation figures reflect to which extent the quality of the vowel in TL condition mimics the individual variation in realization of each stressed vowel.

Table 5. Pearson product-moment correlations for formant values of the stressed vowel and the target vowel for each stressed vowel depending on condition.

Condition	Stressed vowel	F1 correlation	F2 correlation
TL	i	0.515 ($p<0.001$)	0.137 (n.s.)
TL	o	0.347 ($p<0.001$)	0.081 (n.s.)
TVL	i	0.509 ($p<0.001$)	0.365 ($p<0.001$)
TVL	o	0.356 ($p<0.001$)	0.193 ($p<0.01$)

The degree of correlation for F1 is approximately the same in both conditions. However, F2 in the TVL condition seems to reflect the individual variation in the quality of the stressed vowel more than F2 in the TL condition. The gestural sharing hypothesis would predict the opposite since under that assumption the vowel in TL and the preceding vowel arise from the same gesture.

Thus the data on correlation between the quality of stressed vowel and the vowel in TL and TVL does not support the gesture sharing hypothesis. In fact, it was found that TL condition exhibits a weaker correlation for F1 than TVL condition

4. Discussion

Overall, the experimental results show that there is no vowel interval in the word-final LT sequences. The target sequence in the LVT condition has a longer duration and an energy distribution which differs from LT condition in a way consistent with the presence of an additional vowel (Figure 4). On the other hand, the energy distribution in the TL condition is similar to the TVL condition (Figure 5), and the duration difference is small.

The interval between the stop and the liquid in the TL condition is realized differently from the way Russian unstressed schwa is realized: it is possibly slightly shorter, it has lower F1 and F2 but higher F3 (Table 3). In section 4.1 the phonological status of TL sequences is discussed. Section 4.2 relates the results to the concept of transitional vowels. Section 4.3 discusses the role of sonority.

4.1 Phonological status of the vocalic interval in TL

The interval appearing in Russian word-final TL sequences is acoustically different from the unstressed vowel heard in the TVL sequences. The vowel-like sound in TL is quieter and possibly slightly shorter than the lexical vowel, it has lower F1 and F2 but higher F3. Furthermore, the quality of the TL interval does not show significantly higher correlation with the quality of the preceding vowel than the lexical vowel.

If the vocalic interval in TL was a result of phonological epenthesis, we would expect it to have the same quality as Russian unstressed schwa since it occurs in the position of reduction.

This prediction is not borne out by acoustic data. The interval in TL also differs in quality from other Russian vowels occuring in the post-stressed context (Padgett & Tabain 2005). Thus the vowel which would have the acoustic target of the vowel in TL is otherwise unattested in Russian. If we assumed that we are dealing with epenthesis here, it would have to be explained why this vowel is not epenthesized elsewhere and why vowel reduction leads to a different vowel target.

Gouskova & Hall (2010) find that epenthetic /i/ vowels in Lebanese Arabic are produced with lower F2 than lexical vowels by some speakers, while others do not show a difference. Thus in general the qualitative differences themselves do not necessarily show that a vocalic element is not epenthetic. However in the case of Russian the speakers are consistent in producing a qualitatively different vowel in the TL sequence from what they produce in TVL, and the vocalic element in TL differs in all three formanst from TVL. Furthermore, the TL-TVL duration difference is barely perceptible (and possibly arising as a task effect). All of this indicates that in Russian TL clusters, unlike in Lebanese Arabic, the vocalic element cannot be considered as in some way 'in-between' a schwa and a zero. Therefore the line of analysis proposed by Gouskova & Hall (2010) for Lebanese Arabic does not apply to the Russian data in this article.

The acoustic properties of the vocalic interval in TL are better explained on the assumption that it arises from extension and vocalization of the liquid together with an open transition. Specifically, the extension of the vocalic dorsal gesture of /l/ is consistent with the acoustic findings.

The vocalic interval is quieter than a lexical vowel since it does not have its own gestural target. The quality of the vocalic interval is governed by two factors. First, since the vocalic interval arises in a transition from one constricted gesture to another, the mouth is more closed than for the articulation of a lexical vowel – hence the transitional vowel has lower F1 (see Davidson (2006; 2010) for similar findings). Second, the quality of the vocalic interval is influenced by the liquid. Since part of the /l/ target involves a dorsal articulation (Kochetov, 2005; Proctor, 2009, 2011), F2 is lowered. Higher F3 also reflects the influence from /l/ (Stevens, 1998). Finally, the vocalic interval in TL was found to exhibit slightly less coarticulation to the preceding vowel than a lexical vowel. This is also consistent with the assumption that the vocalic interval is part of the realization of /l/ and hence mostly influenced by the consonant.

Some of the findings are also consistent with the hypothesis the vocalic interval in TL constitutes a transitional vowel. However, the high degree of influence from /l/ seen in F2 and F3 makes it unnecessary to postulate a separate transitional vowel. In addition assuming that we are dealing with /l/ extension rather than a separate transitional vowel is supported by the fact that the vocalic interval in TL exhibited slightly less coarticulation with preceding vowel than a lexical vowel.

If the quality of the vocalic interval in /l/-final sequences is influenced most prominently by /l/, it is expected that the /r/-final clusters may exhibit a vocalic interval of a slightly different quality. The exploration of this prediction is left for future research (note also the discussion of the variable realization of Russian /r/ in Iskarous & Kavitskaya (2010)).

Overall, the vocalic interval arising in the TL condition is better interpreted as arising from the extension of the liquid. The vocalic interval does not show higher correlation with the quality of the preceding vowel than the lexical vowel suggesting that it is not arising from the extension of the preceding V gesture as many of the cases studied by Hall (2003; 2006).

4.2 Diagnosing transitional vowels

The realization of Russian word-final TL is very similar to a transitional vowel, and yet the acoustic facts can be better explained if we postulate liquid extension. The Russian case shows that not all vowel-like elements in clusters are transitional vowels. In the case of a sonorant+obstruent cluster there is another possibility to consider: extending the sonorant can produce the vowel-like percept.

Careful acoustic analysis is needed to distinguish between genuine transitional vowels and extended sonorants. In particular, the vowel-like periods arising due to sonorant extension will be mostly influenced by the sonorant whereas the transitional vowel in an open transition may be influenced by both flanking consonants as well as a neighboring vowel. Some studies, such as Schmeyser (2009) also report transitional vowels which have a schwa-like quality disregard of their surroundings.

Even though the transitional vowel hypothesis and the sonorant extension hypothesis sometimes make very similar predictions for the acoustics of the cluster, the two hypotheses have to be kept apart because they make different predictions in other domains.

First, if the coordination patterns in consonant clusters are governed by grammatical constraints in an optimality-theoretic grammar (Gafos 2002), then the presence of a coordination relation necessary for a transitional vowel will imply a constraint ranking which can have effects on other CC-transitions in the language. However, if we are dealing with extension of sonorants within the same grammatical framework, the phenomenon has fewer implications for clusters not containing sonorants.

Second, the appearance of transitional vowels and the variable realization of sonorants may be governed by different principles. Thus Hall (2003; 2006) arguest that transitional vowels are never inserted in response to segmental or syllabic markedness constraints. On the other hand the realization of liquids is known to potentially depend on such factors as syllabic position (Gick *et al.* 2006; Sproat & Fujimura 1993). The Russian data in this article suggest that the realization of Russian liquids depends on the sonority composition of the cluster that the liquid occurs in. The implications of this finding are discussed in section 4.3.

To summarize, the present study implies that transitional vowels should be postulated with caution, especially in clusters of obstruents and sonorants where sonorant extension is an option. Sonorant extension and the transitional vowels can be distinguished based on the acoustic details, and also potentially based on the grammatical factors deterimining their distribution.

4.3 Cluster realization and sonority sequencing

In Russian the extension of liquids only happens in final clusters of rising sonority (TL), but not in LT indicating that the gestural realization of a cluster may depend on its sonority composition. The realization of /l/ may depend on its syllabic position (Gick *et al.* 2006; Sproat & Fujimura 1993). This study shows that not only syllable position but also cluster sonority may play a role in liquid realization. Furthermore, although other sonorants were not tested in the experiment, intuitively the behavior of /r/ and nasals is similar to /l/.

If a word-final TL sequence was realized without an open transition and without liquid extension, the perceptibility of the final liquid could be treatened. Thus a possible reason for liquid extension in this case is that the liquid would not be recoverable otherwise. The acoustic data in this study do not bear on the syllabic status of the extended liquid. Thus the extension may be required because phonologically the liquid becomes syllabic, or it may be a gestural effect of recoverability requirements which does not necessarily create a syllabic consonant (Wright, 1996, 2004; Chitoran *et al.*, 2002).

5. Conclusion

This study examined the acoustics of Russian word-final clusters containing a stop and a liquid. The sequences of rising sonority (TL) are realized with a vowel-like interval while the sequences of falling sonority (LT) are not. The quality and acoustic properties of the vocalic interval are best captured by assuming that it arises via the vocalization of /l/ overlapping the release of the stop.

The quality of the vocalic interval in word-final TL sequences suggests that the interval arises via the extension of a sonorant gesture. The Russian case thus shows that not all vowel-like elements in clusters are transitional vowels.

The correlation with the quality of the preceding vowel is not significantly stronger for the vocalic element in TL than for the lexical vowel. Thus the vocalic element in TL could not arise via gesture sharing with the preceding vowel.

Finally, Russian presents a case where gestural timing within a consonant cluster is dependent on the sonority composition of the cluster.

6. Acknowledgements

I would like to thank Shigeto Kawahara, Paul de Lacy, and Bruce Tesar for their help at all stages of this project. I am also grateful to the audiences at RULING 2011, RUMMIT 2011, NYU PEP Laboratory Meeting, and NELS 2011. For their helpful comments, I am grateful to Adam Albright, Aaron Braver, Vandana Bajaj, Lisa Davidson, Edward Flemming, Gillian Gallagher, Maria Gouskova, Peter Jurgec, John Kingston, Kevin Roon, and Donca Steriade. All errors and imprecisions are my own.

Appendix A. Stimuli and fillers

Condition	Stressed V	Stimulus (cyrillic)	Stimulus (romanized)
LT	i	зафи́лк	zafílk
LT	i	нази́лк	nazílk
LT	i	нати́лк	natílk
LT	i	поди́лк	podílk
LT	i	поси́лк	posílk
LT	o	заро́лк	zarólk
LT	o	захо́лк	zahólk
LT	o	назо́лк	nazólk
LT	o	поко́лк	pokólk
LT	o	поно́лк	ponólk
LVT	i	заги́лок	zagílok
LVT	i	зади́лок	zadílok
LVT	i	нами́лок	namílok
LVT	i	нари́лок	narílok
LVT	i	поки́лок	pokílok

Fillers (cyrillic)	Fillers (romanized)
заби́к	zabík
зари́к	zarík
нахи́к	nahík
пови́к	povík
поти́к	potík
заго́к	zagók
наво́к	navók
назо́к	nazók
поно́к	ponók
пофо́к	pofók
зади́кла	zadíkla
запи́кла	zapíkla
наги́кла	nagíkla
нари́кла	naríkla
поси́кла	posíkla

LVT	o	зафóлок	zafólok
LVT	o	нанóлок	nanólok
LVT	o	натóлок	natólok
LVT	o	погóлок	pogólok
LVT	o	похóлок	pohólok
TL	i	занúкл	zaníkl
TL	i	зарúкл	zaríkl
TL	i	намúкл	namíkl
TL	i	нафúкл	nafíkl
TL	i	покúкл	pokíkl
TL	o	задóкл	zadókl
TL	o	засóкл	zasókl
TL	o	навóкл	navókl
TL	o	позóкл	pozókl
TL	o	попóкл	popókl
TVL	i	зафúкал	zafíkal
TVL	i	надúкал	nadíkal
TVL	i	нанúкал	naníkal
TVL	i	повúкал	povíkal
TVL	i	позúкал	pozíkal
TVL	o	загóкал	zagókal
TVL	o	занóкал	zanókal
TVL	o	напóкал	napókal
TVL	o	нахóкал	nahókal
TVL	o	подóкал	podókal

занóкла	zanókla
надóкла	nadókla
нафóкла	nafókla
побóкла	pobókla
похóкла	pohókla
загúл	zagíl
захúл	zahíl
натúл	natíl
покúл	pokíl
пофúл	pofíl
зарóл	zaról
затóл	zatól
набóл	naból
нафóл	nafól
похóл	pohól
зазúлка	zazílka
зарúлка	zarílka
нанúлка	nanílka
позúлка	pozílka
потúлка	potílka
зафóлка	zafólka
нанóлка	nanólka
нарóлка	narólka
погóлка	pogólka
позóлка	pozólka

Appendix B. Microsoft Word macro used to randomize lines

```
Sub Scramble()
'source: http://www.officekb.com/Uwe/Forum.aspx/ms-
word/4456/Randomize-the-lines
  Dim oTbl As Table
  Dim nRow As Integer, maxRow As Integer

  Application.ScreenUpdating = False

  If (Selection.Type <> wdSelectionNormal) _
     Or (Selection.Paragraphs.Count < 2) Then
     ActiveDocument.Range.Select
```

```
  End If

  Set oTbl = Selection.ConvertToTable(Separator:=vbCr)
  With oTbl
     maxRow = .Rows.Count
     .Columns.Add beforecolumn:=.Columns(1)

     For nRow = 1 To maxRow
        .Cell(nRow, 1).Range.Text = _
           CInt(Rnd() * 10 * maxRow)
     Next nRow

     .Sort excludeheader:=False, _
        fieldnumber:=1, _
        sortfieldtype:=wdSortFieldNumeric, _
        sortorder:=wdSortOrderAscending

     .Columns(1).Delete
     .ConvertToText Separator:=vbCr
  End With

  Application.ScreenUpdating = True
End Sub
```

References

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

Baayen, H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: CUP.

Baayen, H. (2009). *LanguageR*. (R package).

Baayen, H., Davidson, D. & Bates, D. (2008). Mixed-effectsmodeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 300-412.

Barnes, J. (2006a). Phonetics and phonology in Russian unstressed vowel reduction: A study in hyperarticulation. Ms, Boston University.

Barnes, J. (2006b). *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*. Berlin/New York: Mouton de Gruyter.

Bates, D. (2005). Fitting linear mixed models in R. *R news* 5, 27-30.

Bates, D., Maechler, M. & Bolker, B. (2008). *lme4: Linear mixed-effects models using S4 classes*. (R package).

Boersma, P. & Weenink, D. (2011). *Praat: doing phonetics by computer* [online]. Available from: www.praat.org.

Bondarko, L. V. (1977). *Zvukovoi stroi sovremennogo russkogo iazyka*. Moscow: Prosveshchenie.

Bradley, T. (2004). Gestural timing and rhotic variation in Spanish codas. In: Face, T. (Ed.) *Laboratory Approaches to Spanish Phonology*. pp 195-220. Berlin: Mouton de Gruyter.

Bradley, T. (2005). Systemic markedness and phonetic detail in phonology. In: Gess, R. & Rubin, E. (Eds.) *Experimental and Theoretical Approaches to Romance Linguistics*. pp 41-62. Amsterdam: John Benjamins.

Bradley, T. (2007a). Morphological derived-environment effects in gestural coordination: A case study of Norwegian clusters. *Lingua* 117, 950-985.
Bradley, T. (2007b). Spanish complex onsets and the phonetics-phonology interface. In: Colina, S. & Martínez-Gil, F. (Eds.) *Optimality-Theoretical Studies in Spanish Phonology*. pp 15-38. Amsterdam: John Benjamins.
Browman, C. & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook 3* 219-252.
Browman, C. & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica* 45, 140-155.
Chitoran, I., Goldstein, L. & Byrd, D. (2002). Gestural overlap and recoverability: articulatory evidence from Georgian. In: Gussenhoven, C. & Warner, N. (Eds.) *Proceedings of Laboratory Phonology 7*, 2002. Berlin/New York, Mouton de Gruyter.
Colantoni, L. & Steele, J. (2005). Phonetically-driven epenthesis asymmetries in French and Spanish obstruent-liquid clusters. In: Gess, R. & Rubin, E. (Eds.) *Experimental and Theoretical Approaches to Romance Linguistics*. pp 77-96. Amsterdam: John Benjamins.
Davidson, L. (2006). Phonology, phonetics, or frequency: influences on the production of non-native sequences. *Journal of Phonetics* 34, 104-137.
Davidson, L. (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics* 38(2), 272-288.
Davidson, L. & Roon, K. (2008). Durational correlates for differentiating consonant sequences in Russian. *Journal of the International Phonetic Association* 38, 137-165.
Davidson, L. & Stone, M. (2004). Epenthesis versus gestural mistiming in consonant cluster production. In: Tsujimura, M. & Garding, G. (Eds.) *Proceedings of WCCFL22*, 2004. Somerville, MA: Cascadilla Press.
Elson, B. (1956). Sierra Popoluca syllable structure. *International Journal of American Linguistics* 13, 13-17.
Fujisaki, H., Nakamura, K. & Imoto, T. (1973). Auditory perception of duration of speech and non-speech stimuli. *University of Tokyo Research Institute of Logopedics and Phoniatrics Annual Bulletin* 7, 45 - 64.
Gafos, A. (2002). A grammar of gestural coordination. *NLLT* 20, 269-337.
Gick, B., Campbell, F., Oh, S. & Tamburri-Watt, L. (2006). Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. *Journal of Phonetics* 34(1), 49-72.
Gouskova, M. & Hall, N. (2010). Acoustics of epenthetic vowels in Lebanese Arabic. In: Parker, S. (Ed.) *Phonological Argumentation. Essays on Evidence and Motivation*. pp 203-225. London: Equinox.
Hall, N. (2003). *Gestures and Segments: Vowel Intrusion as Overlap*. PhD dissertation, Umass Amherst.
Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. *Phonology* 23, 387-429.
Halle, M. (1959). *The sound pattern of Russian*. The Hague: Mouton.
Iskarous, K. & Kavitskaya, D. (2010). The interaction between contrast, prosody, and coarticulation in structuring phonetic variability. *Journal of Phonetics* 38(4), 625-639.

Jones, D. & Ward, D. (1969). *The phonetics of Russian*. Cambridge: CUP.
Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59(5), 1208-1221.
Kochetov, A. (2005). Phonetic sources of phonological asymmetries: Russian laterals and rhotics. In: Gurski, C. (Ed.) *Proceedings of the Canadian Linguistics Association Annual Conference*, 2005.
Lightner, T. (1965). *Segmental phonology of modern standard Russian*. PhD dissertation, MIT.
Lightner, T. (1972). *Problems in the theory of phonology, Vol. 1: Russian phonology and Turkish phonology*. Edmonton: Linguistic Research Inc.
Matteson, E. & Pike, K. (1958). Non-phonemic transition vocoids in Piro (Arawak). *Miscellanea Phonetica* 3, 22-30.
Padgett, J. (2001). Contrast dispersion and Russian palatalization. In: Hume, E. & Johnson, K. (Eds.) *The role of speech perception in phonology*. pp 187-218. San Diego, CA: Academic Press.
Padgett, J. & Tabain, M. (2005). Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica* 52(1), 14-54.
Proctor, M. (2009). *Gestural characterization of a phonological class: the liquids*. PhD dissertation, Yale University.
Proctor, M. (2011). Towards a gestural characterization of liquids: Evidence from Spanish and Russian. *Laboratory Phonology* 2(2), 451-485.
Purcell, E. (1979). Formant frequency patterns in Russian VCV utterances. *Journal of the Acoustical Society of America* 66(6), 1691-1702.
Quilis, A. (1970). El elemento esvarabático en los grupos [pr, br, tr]. *Phonetique et Linguistique Romaines: Melanges offerts a M. Georges Straka*. pp 99-104. Lyon-Strasbourg: Société de Linguistique Roman.
Quilis, A. (1993). *Tratado de fonología y fonética españolas*. Madrid: Editorial Gredos.
R Development Core Team (2011). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
Ramírez, C. J. (2002). Characterization of the epenthetic vowel between the clusters formed by stops/fricatives + flap in Spanish. In: Spreng, B. (Ed.) *Proceedings of the Annual Meeting of the Niagara Linguistic Society 19*, 2002. pp 67-74. Toronto: University of Toronto Linguistics Graduate Course Union.
Ramírez, C. J. (2006). Acoustic and Perceptual Characterization of the Epenthetic Vowel between the Clusters Formed by Consonant + Liquid in Spanish. In: Díaz-Campos, M. (Ed.) *Proceedings of the Second Conference on Laboratory Approaches to Spanish Phonetics and Phonology,* 2006. pp 48-61. Somerville, MA: Cascadilla Proceedings Project.
Scherba, L. V. (1912). Russkie glasnye v kachestvennom i kolichestvennom otnoshenii. Ms, Saint Petersburg.
Schmeiser, B. (2009). An acoustic analysis of intrusive vowels in Guatemalan Spanish /rC/ clusters., 2009.
Sharoff, S. (2002). The frequency dictionary for Russian. Available from: http://bokrcorpora.narod.ru/frqlist/frqlist-en.html.
Shaw, J. & Davidson, L. (2011). Perceptual similarity in input–output mappings: a

computational/experimental study of non-native speech production. *Lingua* 121, 1344-1358.
Sproat, R. & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21, 291-311.
Steriade, D. (1990). Gestures and autosegments: comments on Browman and Goldstein's paper. In: Kingston, J. & Beckman, M. (Eds.) *Papers in Laboratory Phonology. I: Between the grammar and physics of speech*. pp 382-297. Cambridge, Mass: CUP.
Stevens, K. (1998). *Acoustic Phonetics*. Cambridge, Mass: MIT Press. (Current studies in linguistics; 30).
Timberlake, A. (2004). *A reference grammar of Russian*. Cambridge: CUP.
Turk, A., Nakai, S. & Sugahara, M. (2006). Acoustic segment durations in prosodic research: a practical guide. In: Sudhoff, S., Lenertova, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N., & Schließer, J. (Eds.) *Methods in Empirical Prosody Research*. pp 1-28. Berlin/New York: de Gruyter.
Ward, D. (1975). Unaccented vowels in Russian. *Russian Linguistics* 12(1/2), 91-104.
Wright, R. (1996). *Consonant clusters and cue preservation in Tsou*. PhD dissertation, UCLA.
Wright, R. (2004). Areview of perceptual cues and cue robustness. In: Hayes, B., Kirchner, R., & Steriade, D. (Eds.) *Phonetically based phonology*. pp 34-57. Cambridge: CUP.
Zsiga, E. (2000). Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28, 69-102.
Zsiga, E. (2003). Articulatory timing in a second language. Evidence from Russian and English. *Studies in Second Language Acquisition* 25, 399-432.