Sonorant Extension, Transitional Vowels, and Gesture Sharing: an Acoustic Study of Russian Clusters*

Peter Staroverov

Rutgers, the State University of New Jersey

1. Introduction

The study of cluster realizaition within the framework of articulatory phonology (Browman and Goldstein, 1986, 1988 et seq.) has uncovered a number of cases where the cluster is realized with a vocalic interval which nonetheless does not appear to be part of the phonological representation (see Bradley, 2004, 2007; Davidson, 2006, 2010; Gafos, 2002; Hall, 2003, 2006 among others). Based on acoustic data from Russian word-final clusters, this paper analyzes possible gestural configurations which may lead to appearance of these vocalic intervals. In addition, the effects of the ordering of consonants within the cluster, and of concomitant changes in sonority sequencing, are addressed.

An acoustic study of the realization of Russian word-final clusters of stops and liquids is presented. A vocalic interval was found in the stop-liquid clusters (as in /tsikl/ 'cycle'). The acoustic properties of this interval indicate that it arises via extension of the liquid gesture, rather than the default voicing state of the glottis. It is argued that in the case of Russian, the perceived vocalic interval is better analyzed as part of the liquid realization, rather than a separate transitional vowel.

The realization of liquid-stop clusters (as in /polk/ 'regiment') did not exhibit a vocalic interval. Thus the sonority composition of a cluster seems to play a role in how this cluster is realized. It is hypothesized that the effect of sonority has to do with the fact that the liquid is less perceptible in the final stop-liquid clusters.

1.1 Cluster Realizaton and Transitional Vowels

Even when there is no vowel target between two consonants, there are a number of

* 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.

articulatory configurations which may lead to a percept of a vowel within a consonant cluster. In all of these cases there is a lag between the time when the constriction for the first consonant is released and the time when the second consonant reaches its target. This temporal lag is a period of relatively unobstructed vocal truct which leaves some space for an acoustic vowel to appear in.

There are a number of reasons why a vocalic element may arise in a temporal lag between two consonant constrictions. First, Browman and Goldstein (1986) suggest that the default state of the glottis is one that produces voicing. Thus if the two consonant constrictions are relaxed for long enough, a period of periodicity can occur. The quality of this vocalic interval is somewhat similar to that of a schwa, although the mouth is in a more constricted position than for the articulation of a lexical schwa (Davidson, 2006, 2010).

When one of the consonants in the cluster is a sonorant the realization of the consonant itself may potentially yield a vocalic interval in an open transition. In these cases the sonorant is extended to overlap the open transition between the two consonants, and thus there may be no need to postulate a separate transitional vowel. The quality of such vocalic intervals will be mostly affected by the sonorant. This paper argues that Russian word-final stop-liquid clusters involve such a configuration.

Finally, Steriade (1990) proposes that some 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.

Hall argues that obstruent consonant gestures are very rarely overlapped by neighboring vowels. The sequences to be considered in this article are of the form vowel-stop-liquid: VTL#. For an echo vowel to arise in such sequences the lexical vowel would have to completely overlap the stop gesture, and this is unlikely according to Hall's survey. The hypothesis that the vocalic interval in Russian TL# clusters arises because of gesture sharing with a neighboring vowel was considered in this study, but it was not supported by the acoustic data. Since this hypothesis is unlikely to be true given what we know from typology, it will not be discussed in what follows due to space limitations.

To summarize, the vocalic intervals in consonant clusters are assumed to arise due to a relatively open transition. However, there may be different gestural sources of the vowel occuring in this transition, and the exact source of the vowel will determine its quality. Unfortunately the quality of the gestural mistiming vowels is rarely systematically studied. 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 its F1 was lower than for a lexical schwa in both languages.

This study identifies a vocalic interval in Russian TL# clusters. The examination of formant structure indicates that this interval is highly affected by the liquid, and thus most likely occuring due to liquid extension. Thus a close look at the vocalic quality may be needed to tell apart the transitional vowels and the extended sonorants. The Russian data also indirectly confirm that gesture sharing is unlikely across an obstruent.

1.2 Sequential Constraints on Cluster realization

The relation between consonant gestures comprising the cluster affects the degree of gestural overlap which is needed to produce a transitional vowel. Gafos (2002) finds that the same degree of overlap may produce a vocalic interval in heterorganic clusters but no such interval in homorganic clusters since there is no movement from one constriction to the other in the latter case.

The sonority composition of a consonant cluster may also affect its gestural 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. In addition, the liquid /l/ is known to often be realized differently depending on its syllabic position (Sproat and Fujimura, 1993; Gick *et al.*, 2006).

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. If the liquids occur on their own or precede a consonant, the extension is not observed. Furthermore the TL combinations do not exhibit liquid extension when they occur intervocalically[1]. It is hypothesized that the perceptibility of the liquid is particularly threatened in the coda TL clusters compared to the cases when the liquid occurs prevocalically or postvocalically. Therefore the liquid gets extended just in those cases.

1.3 Russian Consonant Clusters

Russian liquids contrast in palatalization: /l/ - /l^j/, /r/ - /r^j/. This study focuses on nonpalatalized liquids. According to native speaker intuitions, palatalized and nonpalatalized liquids behave the same way in consonant clusters. In an acoustic study of Russian rhotics Iskarous and 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 (Zsiga, 2000, 2003; Davidson and Roon, 2008). 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 and 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-

[1] Thus Russian TL clusters are not susceptible to vowel intrusion across the board (Flemming, 2008). I am grateful to Donca Steriade for discussing this hypothesis with me.

final clusters of rising sonority (1)a intuitively seem to be pronounced with vocalization. In Standard Russian, /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 while /i/ and /e/ neutralize to [ɪ] in both of these positions (see e.g. Barnes (2006a, 2006b) and references therein). In (1) and throughout the paper stress is marked in the words which have more than one syllable.

(1) Russian word-final combinations of a stop and a liquid
  a. T + L
    /bobr/ ‘beaver’; /rublʲ/ ‘rouble’;
    /tʲeátr/ ‘theater’; /podl/ ‘mean.PRED’
  b. T + L
    /polk/ 'regiment' /kolʲt/ 'colt'
    /dolg/ 'debt, duty' /skarb/ 'stuff, goods and chattels'

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 (2) is truly rare.

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

The words which have a word-final TL are less frequent than words with final clusters of falling sonority (Proctor, 2009). Many of the TL words are borrowings, some of them fully assimilated. However the grammatical knowledge of Russian has to include the knowledge of how the word-final TL clusters are pronounced. There are native words ending in TL clusters. The combined token frequency of all tokens with word-final stop-liquid sequences is 290.13 items per million (Sharoff, 2002). Thus these words are heard with some frequency in Russian speech. Furthermore, the formation of genitive plural and predicative forms of adjectives regularly yields word-final TL sequences.

Word-medial stop-liquid coda sequences occur only in a handful of words, most notably those with the prefix /kontr-/ 'counter-' (e.g. /kontrrazvʲédka/ 'counter-intelligence') and the derivatives of the root /bodr/ 'cheerful' (e.g. /bódrstvovatʲ/ 'be awake').

To sum up, intuitively the word-final TL sequences in Russian are realized with a phonetic vowel-like interval whereas LT sequences are not. The experiment in this article confirms these intuitions and addresses the status of the vocalic interval heard in the TL sequences.

2. Method

2.1 Stimuli

The experiment compares four kinds of underlying word-final sequences in a nonce-word production task: TL#, TVL#, LT#, LVT#. In order to address whether there is a vocalic interval in TL and in LT sequences, the difference between TL and TVL is compared to the LT vs. LVT difference.

The comparison between TL and TVL condition addresses the status of the vowel in TL (if present). First, the vowel could result from phonological epenthesis. This hypothesis gets some support if the vowel in TL is realized the same as a lexical vowel. 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 result of gestural mistiming or extension - the gestural sources discussed in section 1.1.

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.

/k/ was chosen for stop quality in order to maximize the word-likeliness of the stimuli. Since Russian post-tonic vowel reduction leads to complete neutralization of /o/ and /a/ (Barnes 2006a, b; Padgett and Tabain 2005) the vowel in /-kal/ words was the same as in /-lok/ words. The orthographic vowel was varied to increase the word-likeliness. The similarity to existing words was assessed based on Sharoff's (2002) frequency dictionary. Finally, /l/ was chosen as the target liquid because the realization of the other Russian liquid /r/ varies greatly across tokens (Iskarous and Kavitskaya, 2010).

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. The quality of the second syllable vowel was varied in the experiment, and the preceding consonant varied accordingly. This variation was used to address the gesture sharing hypothesis. Table 1 summarizes the experimental conditions and gives sample stimuli for each condition.

Type of sequence	**Final sequence**	**Transliterated examples**	**Transcribed examples**
-TL	-okl/-ikl	zadókl, nafʲíkl	[zɐdókᵊl], [nɐfʲíkᵊl]
-TVL	-okəl/-ikəl	zagókal, pozʲíkal	[zɐgókəl], [pɐzʲíkəl]
-LT	-olk/-ilk	nazólk, posʲílk	[nɐzólk], [pɐsʲílk]
-LVT	-olək/-ilək	nanólok, pokʲílok	[nɐnólək], [pɐkʲílək]

Table 1. *Experimental conditions and sample stimuli*

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. Just like the stimuli, 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" and started with "na", "za" or "po".

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 (3) and presented to the speakers in Russian orthography with stress marked on the target word. Within this sentence the target word received focus since it contained new information.

(3)	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 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 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 and Weenink 1999-2011). 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).

2.5 Acoustic analysis and statistics

All acoustic measurements were taken using Praat (Boersma and 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.

In the TL and TVL condition the target interval included both the liquid and the vowel after the stop. The formants representative of the vocalic part were measured at the end of the third window (i. e. the 3/20 of the interval duration). 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.

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 excluded from formant measurements.

The statistical analysis was done in R (R Development Core Team, 2011). 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.

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]/[ol] in LT stimuli and [ilə]/[olə] in LVT stimuli). 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); $p<0.001$.

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. Thus the [l] and the preceding vocalic interval was included. 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. A negative difference indicates that TL interval was longer on average. The significance of the differences was assessed using a two-sided non-paired t-test. The significance values in **Table 2** reflect Bonferroni adjustment. Awareness of the TL vs. TVL contrast is shown based on post-experiment feedback.

Speakers	**TL - TVL range (ms)**	**t.test p-value (Bonf. adjusted)**	**Aware of the contrast**
EK	-8.6	n.s.	no
ES, OE, KG, DS	- 2.5 to +2.5	n.s.	no
BN, PC, MT, KR, EP	> +10	p<0.005	4 out of 5

Table 2. *Duration differences between the target interval in TVL and TL condition by speaker*

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 and one speaker showed a reversal – the vowel interval in the TL condition was longer by 8.6 milliseconds (n.s.).

These results suggest that 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 vocalic interval

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. The target interval was broken down into 20 equal windows and mean intensity within each window was calculated. **Figure 1** shows the schematized distribution of energy across conditions.

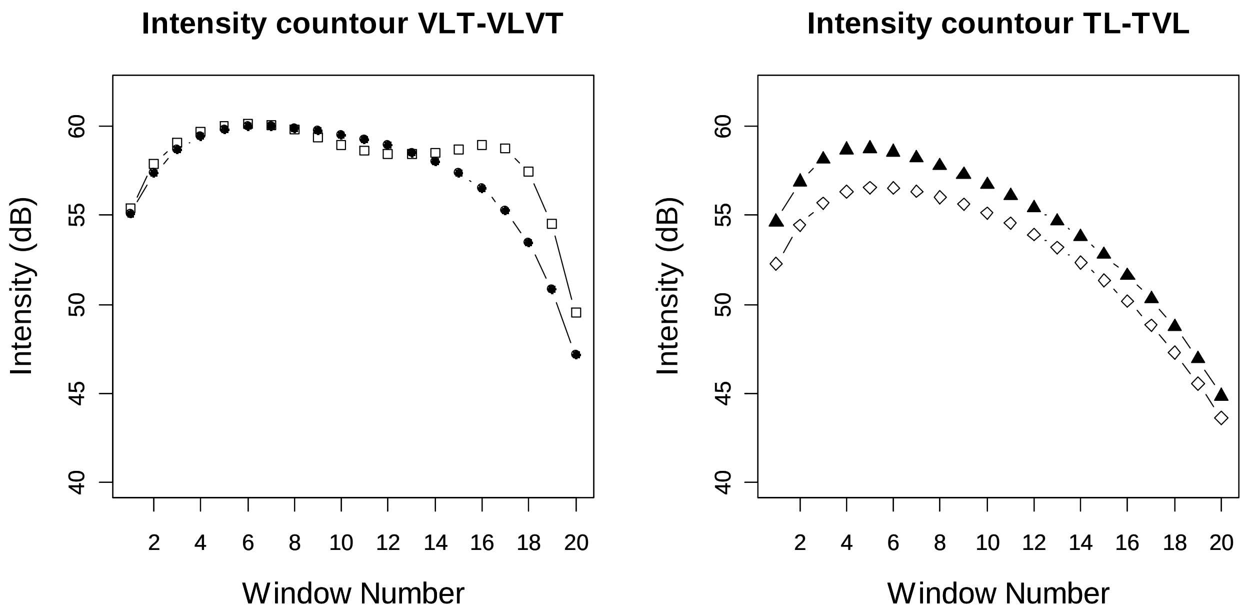

Figure 1. *Intensity contours across conditions: (V)LT (dark bullets), (V)LVT (squares), TL (diamonds), TVL (dark triangles)*

Considering the graph for VLT and VLVT conditions (**Figure 1**, left panel) it should be kept in mind that the two lines of the graph correspond to different time periods since the duration of the target interval was longer in VLVT. With this complication in mind, the energy distribution in VLT and VLVT is different. Both lines show a comparable rise of intensity for the stressed vowel. The graph for the VLVT condition has a slight rise in intensity after the first peak corresponding to the realization of the second underlying vowel. However the graph for the VLT condition shows a gradual fall after the first peak showing no vocalic interval between the liquid and the stop.

The energy distribution in the TL and TVL conditions is shown in the right panel of **Figure 1**. In this case the target interval excluded the strssed vowel. The two intensity contours are similar. There is one intensity peak suggesting that a vocalic interval after the stop 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 of considerable intensity (windows 4-7). The loudness of the signal in that period is on the range of the intensity of the vowel in the first syllable of the carrier phrase.

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

3.3 Acoustics of the vowel in TL

In order to assess the status of the vocalic interval in TL condition, its duration, intensity and formant structure was compared to the TVL condition. 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 (**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 1**). It was found that intensity was lower in the TL condition on average: 56.3 dB (±0.52dB) vs. 58.7 dB (±0.49dB) ($p<0.001$).

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.

	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**

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

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. Thus there was some V-to-V coarticulation with the preceding vowel. However, a closer look at the correlation in vowel quality showed that the data do not line up with what we would expect if the two vowels were the same gesture.

In general, the formant values for each speaker followed the 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.

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

4. Discussion

Overall, the experimental results show that there is no vowel interval in the word-final LT sequences, but there is a vocalic interval in the TL condition. The interval between the stop and the liquid in the TL condition is realized differently from the way Russian unstressed schwa is realized. In section 4.1 possible sources of the vocalic interval in TL are discussed. Section 4.2 discusses the role of sonority in cluster realization.

4.1 The Sources of the Vocalic Interval and the Typology of Transitional Vowels

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.

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. In some languages the epenthetic vowels are reported to be qualitatively 'in-between' the phonological target and zero (Gouskova and Hall, 2010). The case of Lebanese Arabic reported by Gouskova and Hall involves between-speaker variation in vowel quality, so that some speakers produce more reduction than others. However in the case of Russian the speakers are fairly consistent in producing a qualitatively different vowel in the TL sequence from what they produce in TVL. Thus unlike in Lebanese Arabic, the vocalic interval in Russian TL clusters does not appear to be gradually closer to zero than the vowel in TVL, it just has a different quality. Hence the vocalic interval in Russian does not seem to correspond to a reduced epenthesized vowel.

If the realization of TL clusters involves an open transition, there is a number of possible gestural sources of the vocalic interval (see 1.1). First, the vocalic interval could arise because of the default voicing state of the vocal tract, in which case it would be realized as a neutral vowel, possibly more closed than a schwa (Davidson, 2006, 2010). Second, liquid extension could lead to a percept of a period of voicing during the open transition. In this latter case, the vocalic interval is expected to have a quality of a highly l-colored schwa.

The extension of the vocalic dorsal gesture of /l/ provides an explanation for the acoustic properties of the interval. The vocalic interval is quieter than a lexical vowel since it does not have its own gestural target. 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 it has lower F1 (Davidson, 2006, 2010).

The low F1 of the vocalic interval does not differentiate between the two gestural hypotheses outlined above. However the F2 and F3 are highly influenced by /l/. Since part of the /l/ target involves a dorsal articulation (Kochetov, 2005; Proctor, 2009, 2011), F2 is lower than that of a schwa. Higher F3 also reflects the influence from /l/ (Stevens, 1998). These formant differences between the TL vowel and schwa would be unexpected if we assumed that we are dealing with some sort of a default transitional vowel.

Thus the vocalic interval in Russian word-final TL clusters most likely is due to the extension of the liquid gesture. This makes it unnecessary to postulate a separate transitional 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.

In general, this study shows that not all vowel-like elements in clusters are transitional vowels. In the case of a sonorant+obstruent cluster there is an additional 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. The vowel-like periods arising due to sonorant extension will be mostly influenced by the sonorant. On the other hand, transitional vowel stemming from the default voicing in the glottis would tend to have a schwa-like quality.

Even though the acoustics of clusters involving transitional vowels is somewhat similar to clusters with sonorant extension, the two hypotheses have to be kept apart because the appearance of transitional vowels and the variable realization of sonorants may be governed by different principles. 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 and Fujimura, 1993).

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 transitional vowels can be distinguished based on the acoustic record and also potentially based on the grammatical factors deterimining their distribution.

4.2 Cluster Realization and Sonority Sequencing

In Russian, the extension of liquids only happens in final clusters of rising sonority (stop-liquid), but not in liquid-stop. This suggests 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/.

A realization of a word-final stop-liquid sequence without an open transition and without liquid extension could treaten the perceptibility of the final liquid. Thus a possible reason for liquid extension in this case is that it would not be recoverable otherwise.

The acoustic data in this study do not bear on the relation between liquid perceptibility and its phonological status. It is possible that liquid extension is required because phonologically the liquid becomes syllabic. On the other hand, liquid extension may be a direct gestural effect of recoverability requirements (Wright, 1996, 2004; Chitoran *et al.*, 2002).

5. Conclusion

Based on acoustic data from a nonce-word production task it was shown that Russian word-final stop-liquid sequences contain a vocalic interval. The quality of the vocalic interval is best explained on the assumption that it arises as a by-product of extending the liquid gesture to overlap the open transition between two consonants. Thus sonorant extension in consonant clusters yields an acoustic result similar to a transitional vowel.

Detailed analysis of vowel quality is needed to differentiate between extended sonorants and genuine transitional vowels.

No vocalic interval was found in the word-final liquid-stop clusters. This indicates that the sonority composition of a cluster may influence how the cluster is realized.

References

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

Barnes, J., 2006b. Phonetics and phonology in Russian unstressed vowel reduction: A study in hyperarticulation.

Boersma, P. and Weenink, D., 2011. *Praat: doing phonetics by computer*. Computer program.

Bradley, T., 2004. Gestural timing and rhotic variation in Spanish codas. *In*: T. Face, ed. *Laboratory Approaches to Spanish Phonology*. Berlin: Mouton de Gruyter, 195–220.

Bradley, T., 2007. Morphological derived-environment effects in gestural coordination: A case study of Norwegian clusters. *Lingua*, 117, 950–985.

Browman, C. and Goldstein, L., 1986. Towards an articulatory phonology. *Phonology Yearbook 3*, 219–252.

Browman, C. and Goldstein, L., 1988. Some notes on syllable structure in articulatory phonology. *Phonetica*, 45, 140–155.

Chitoran, I., Goldstein, L., and Byrd, D., 2002. Gestural overlap and recoverability: articulatory evidence from Georgian. *In*: C. Gussenhoven and N. Warner, eds. *Laboratory Phonology 7*. Berlin/New York, Mouton de Gruyter.

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. and Roon, K., 2008. Durational correlates for differentiating consonant sequences in Russian. *Journal of the International Phonetic Association*, 38, 137–165.

Flemming, E., 2008. Asymmetries between assimilation and epenthesis. Ms., MIT

Fujisaki, H., Nakamura, K., and 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., and 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. and Hall, N., 2010. Acoustics of epenthetic vowels in Lebanese Arabic. *In*: S. Parker, ed. *Phonological Argumentation. Essays on Evidence and Motivation*. London: Equinox, 203–225.

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.

Iskarous, K. and Kavitskaya, D., 2010. The interaction between contrast, prosody, and coarticulation in structuring phonetic variability. *Journal of Phonetics*, 38 (4), 625–639.

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*: C. Gurski, ed. *Proceedings of the 2005 Canadian Linguistics Association Annual Conference*.

Padgett, J. and 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.

R Development Core Team, 2011. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Sharoff, S., 2002. The frequency dictionary for Russian. Available at http://www.comp.leeds.ac.uk/ssharoff/

Sproat, R. and 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*: J. Kingston and M. Beckman, eds. *Papers in Laboratory Phonology*. *I: Between the grammar and physics of speech*. Cambridge, Mass: CUP, 382–297.

Stevens, K., 1998. *Acoustic Phonetics*. Cambridge, Mass: MIT Press.

Turk, A., Nakai, S., and Sugahara, M., 2006. Acoustic segment durations in prosodic research: a practical guide. *In*: S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, and J. Schließer, eds. *Methods in Empirical Prosody Research*. Berlin/New York: de Gruyter, 1–28.

Wright, R., 1996. *Consonant clusters and cue preservation in Tsou*. PhD dissertation, UCLA.

Wright, R., 2004. A review of perceptual cues and cue robustness. *In*: B. Hayes, R. Kirchner, and D. Steriade, eds. *Phonetically based phonology*. Cambridge: CUP, 34–57.

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.

Department of Linguistics
Rutgers University
18 Seminary Place
New Brunswick, NJ 08901

petesta@rutgers.edu