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 S1 to coincide with landmark B of the segment S2.

(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)

![Gestural landmarks diagram](attachment:image.png)

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. ![Example a](attachment:image1.png)

b. ![Example b](attachment:image2.png)

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)).
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 preceding 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.

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' /
   /dolg/ 'debt, duty' /

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.
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. /kontrrazvʲédka/ 'counter-intelligence', /kontrrʲevolʲutsiónnij/ 'counter-revolutionary') and the derivatives of the root /bodr/ 'cheerful' (e.g. /bodrstvovatʲ/ '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.

/i/ 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

<table>
<thead>
<tr>
<th>Type of sequence</th>
<th>Preceding V</th>
<th>Final sequence</th>
<th>Transliterated example</th>
<th>Transcribed example</th>
</tr>
</thead>
<tbody>
<tr>
<td>-TL</td>
<td>o</td>
<td>-okl</td>
<td>zadókl</td>
<td>[zedókl'í]</td>
</tr>
<tr>
<td></td>
<td>(i)</td>
<td>-ikl</td>
<td>naříkl</td>
<td>[naříkl]</td>
</tr>
<tr>
<td>-TVL</td>
<td>o</td>
<td>-okl</td>
<td>zagoál</td>
<td>[zagoál]</td>
</tr>
<tr>
<td></td>
<td>(i)</td>
<td>-ikl</td>
<td>pożíkal</td>
<td>[požíkal]</td>
</tr>
<tr>
<td>-LT</td>
<td>o</td>
<td>-olk</td>
<td>nazólk</td>
<td>[nazólk]</td>
</tr>
<tr>
<td></td>
<td>(i)</td>
<td>-ilk</td>
<td>pośílk</td>
<td>[pośílk]</td>
</tr>
<tr>
<td>-LVT</td>
<td>o</td>
<td>-olk</td>
<td>nanólk</td>
<td>[nanólk]</td>
</tr>
<tr>
<td></td>
<td>(i)</td>
<td>-ilk</td>
<td>pokílok</td>
<td>[pokílok]</td>
</tr>
</tbody>
</table>

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. Within this sentence the target word received focus since it contained new information.

\[(7)\]  
\[nɐjdʲɪ'tʲɪ slóvə X f slavrʲé \]
\[\text{find.IMP.PL word.ACC X in dictionary.LOC} \]
\[\text{'Find the word X in the dictionary'} \]

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 VIV 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 /pokʲikl/ by speaker MT. The vowel in
parentheses in the transcription corresponds to the hypothesized vocoid in /kl/.
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 ([æj] in the beginning of the phrase). Within the carrier phrase the baseline interval was followed by a /d/ which was occasionally 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 consideration 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 occasionally 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 [I] in the production of /natʲilk/ and [Iə] 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).

![Waveform images](image.png)

Figure 3. Target interval plus the preceding stop from speaker MT’s third productions of /nadɪkal/ (right panel) and /nafɪkl/ (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.

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

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).
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).

![Normalized Intensity TL and TVL](image)
Table 3. Formant differences between TL and TVL condition. Negative difference means that the relevant formant is lower in the TL condition.

<table>
<thead>
<tr>
<th></th>
<th>F1 (Hz)</th>
<th>F2 (Hz)</th>
<th>F3 (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TL condition</td>
<td>423 (±5)</td>
<td>896 (±14)</td>
<td>2751 (±35)</td>
</tr>
<tr>
<td>TVL condition</td>
<td>486 (±8)</td>
<td>1006 (±14)</td>
<td>2670 (±35)</td>
</tr>
<tr>
<td>Difference and p-value</td>
<td>-63, p&lt;0.001</td>
<td>-110, p&lt;0.001</td>
<td>81, p &lt; 0.001</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>F1 in TL condition</th>
<th>F1 in TVL condition</th>
<th>F2 in TL condition</th>
<th>F2 in TVL condition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>r</strong></td>
<td>0.19 (p&lt;0.001)</td>
<td>0.33 (p&lt;0.001)</td>
<td>0.47 (p&lt;0.001)</td>
<td>0.46 (p&lt;0.001)</td>
</tr>
</tbody>
</table>

3 In a study of vowel quality in Russian VCV sequences Purcell (1979) only finds evidence of V1-V2 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.

Figure 6. Normalized formant correlation scatterplots in the TL and TVL conditions.
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-transformation 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>
<thead>
<tr>
<th>Condition</th>
<th>Stressed vowel</th>
<th>F1 correlation</th>
<th>F2 correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TL i</td>
<td>0.515 (p&lt;0.001)</td>
<td>0.137 (n.s.)</td>
<td></td>
</tr>
<tr>
<td>TL o</td>
<td>0.347 (p&lt;0.001)</td>
<td>0.081 (n.s.)</td>
<td></td>
</tr>
<tr>
<td>TVL i</td>
<td>0.509 (p&lt;0.001)</td>
<td>0.365 (p&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td>TVL o</td>
<td>0.356 (p&lt;0.001)</td>
<td>0.193 (p&lt;0.01)</td>
<td></td>
</tr>
</tbody>
</table>

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 occurring 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 formant values from TVL. Furthermore, the TL-TV 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 determining 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 treated. 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

<table>
<thead>
<tr>
<th>Condition</th>
<th>Stressed V</th>
<th>Stimulus (cyrillic)</th>
<th>Stimulus (romanized)</th>
<th>Fillers (cyrillic)</th>
<th>Fillers (romanized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LT</td>
<td>i</td>
<td>зафи́лк</td>
<td>zafílk</td>
<td>заби́к</td>
<td>zabík</td>
</tr>
<tr>
<td>LT</td>
<td>i</td>
<td>нази́лк</td>
<td>nazílk</td>
<td>зари́к</td>
<td>zarík</td>
</tr>
<tr>
<td>LT</td>
<td>i</td>
<td>нати́лк</td>
<td>natílk</td>
<td>нахи́к</td>
<td>nahík</td>
</tr>
<tr>
<td>LT</td>
<td>i</td>
<td>поди́лк</td>
<td>podílk</td>
<td>пови́к</td>
<td>povík</td>
</tr>
<tr>
<td>LT</td>
<td>i</td>
<td>поси́лк</td>
<td>posílk</td>
<td>поти́к</td>
<td>potík</td>
</tr>
<tr>
<td>LT</td>
<td>o</td>
<td>заро́лк</td>
<td>zarólk</td>
<td>заго́к</td>
<td>zagók</td>
</tr>
<tr>
<td>LT</td>
<td>o</td>
<td>захо́лк</td>
<td>zahólk</td>
<td>наво́к</td>
<td>navók</td>
</tr>
<tr>
<td>LT</td>
<td>o</td>
<td>назо́лк</td>
<td>nazólk</td>
<td>назо́к</td>
<td>nazók</td>
</tr>
<tr>
<td>LT</td>
<td>o</td>
<td>поко́лк</td>
<td>pokólk</td>
<td>поно́к</td>
<td>ponók</td>
</tr>
<tr>
<td>LT</td>
<td>o</td>
<td>поно́лк</td>
<td>ponólk</td>
<td>пофо́к</td>
<td>pofók</td>
</tr>
<tr>
<td>LVT</td>
<td>i</td>
<td>заги́лок</td>
<td>zagílok</td>
<td>зади́кла</td>
<td>zadíkla</td>
</tr>
<tr>
<td>LVT</td>
<td>i</td>
<td>зади́лок</td>
<td>zadílok</td>
<td>запи́кла</td>
<td>zapíkla</td>
</tr>
<tr>
<td>LVT</td>
<td>i</td>
<td>нами́лок</td>
<td>namílok</td>
<td>наги́кла</td>
<td>nagíkla</td>
</tr>
<tr>
<td>LVT</td>
<td>i</td>
<td>нари́лок</td>
<td>narílok</td>
<td>нари́кла</td>
<td>naríkla</td>
</tr>
<tr>
<td>LVT</td>
<td>i</td>
<td>поки́лок</td>
<td>pokílok</td>
<td>поси́кла</td>
<td>posíkla</td>
</tr>
<tr>
<td>LVT</td>
<td>o</td>
<td>зафо́лок</td>
<td>zafо́lok</td>
<td>зано́кла</td>
<td>zanо́kla</td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>LVT</td>
<td>o</td>
<td>нано́лок</td>
<td>nanо́lok</td>
<td>надо́кла</td>
<td>nadо́kla</td>
</tr>
<tr>
<td>LVT</td>
<td>o</td>
<td>нато́лок</td>
<td>natо́lok</td>
<td>нафо́кла</td>
<td>nafо́kla</td>
</tr>
<tr>
<td>LVT</td>
<td>o</td>
<td>пого́лок</td>
<td>pogо́lok</td>
<td>побо́кла</td>
<td>pobо́kla</td>
</tr>
<tr>
<td>LVT</td>
<td>o</td>
<td>похо́лок</td>
<td>похо́лок</td>
<td>похо́кла</td>
<td>похо́кла</td>
</tr>
<tr>
<td>TL</td>
<td>i</td>
<td>заны́кл</td>
<td>zanы́kl</td>
<td>загы́л</td>
<td>zagы́l</td>
</tr>
<tr>
<td>TL</td>
<td>i</td>
<td>зары́кл</td>
<td>zarы́kl</td>
<td>захы́л</td>
<td>zahы́л</td>
</tr>
<tr>
<td>TL</td>
<td>i</td>
<td>нами́кл</td>
<td>намы́кл</td>
<td>наты́л</td>
<td>натыми́л</td>
</tr>
<tr>
<td>TL</td>
<td>i</td>
<td>нафы́кл</td>
<td>нафы́кл</td>
<td>покы́л</td>
<td>покы́л</td>
</tr>
<tr>
<td>TL</td>
<td>i</td>
<td>покы́кл</td>
<td>покы́кл</td>
<td>пофи́л</td>
<td>пофи́л</td>
</tr>
<tr>
<td>TL</td>
<td>o</td>
<td>зады́кл</td>
<td>зады́кл</td>
<td>зары́л</td>
<td>зары́л</td>
</tr>
<tr>
<td>TL</td>
<td>o</td>
<td>засы́кл</td>
<td>засы́кл</td>
<td>зы́л</td>
<td>зы́л</td>
</tr>
<tr>
<td>TL</td>
<td>o</td>
<td>навы́кл</td>
<td>навы́кл</td>
<td>ны́л</td>
<td>ны́л</td>
</tr>
<tr>
<td>TL</td>
<td>o</td>
<td>pozы́кл</td>
<td>pozы́кл</td>
<td>нафы́л</td>
<td>нафы́л</td>
</tr>
<tr>
<td>TL</td>
<td>o</td>
<td>попы́кл</td>
<td>попы́кл</td>
<td>попы́л</td>
<td>попы́л</td>
</tr>
<tr>
<td>TVL</td>
<td>i</td>
<td>зафи́кал</td>
<td>зафы́кал</td>
<td>зафи́кал</td>
<td>зафи́кал</td>
</tr>
<tr>
<td>TVL</td>
<td>i</td>
<td>нады́кал</td>
<td>нады́кал</td>
<td>нады́кал</td>
<td>нады́кал</td>
</tr>
<tr>
<td>TVL</td>
<td>i</td>
<td>нами́кал</td>
<td>нами́кал</td>
<td>нами́кал</td>
<td>нами́кал</td>
</tr>
<tr>
<td>TVL</td>
<td>i</td>
<td>повы́кал</td>
<td>повы́кал</td>
<td>повы́кал</td>
<td>повы́кал</td>
</tr>
<tr>
<td>TVL</td>
<td>i</td>
<td>позы́кал</td>
<td>позы́кал</td>
<td>позы́кал</td>
<td>позы́кал</td>
</tr>
<tr>
<td>TVL</td>
<td>o</td>
<td>загы́кал</td>
<td>загы́кал</td>
<td>загы́кал</td>
<td>загы́кал</td>
</tr>
<tr>
<td>TVL</td>
<td>o</td>
<td>заны́кал</td>
<td>заны́кал</td>
<td>заны́кал</td>
<td>заны́кал</td>
</tr>
<tr>
<td>TVL</td>
<td>o</td>
<td>напы́кал</td>
<td>напы́кал</td>
<td>напы́кал</td>
<td>напы́кал</td>
</tr>
<tr>
<td>TVL</td>
<td>o</td>
<td>нахо́кал</td>
<td>нахо́кал</td>
<td>нахо́кал</td>
<td>нахо́кал</td>
</tr>
<tr>
<td>TVL</td>
<td>o</td>
<td>поды́кал</td>
<td>поды́кал</td>
<td>поды́кал</td>
<td>поды́кал</td>
</tr>
</tbody>
</table>

**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
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


Scherba, L. V. (1912). Russkie glasnye v kachestvennom i kolichestvennom otmoshenii. Ms, Saint Petersburg.


computational/experimental study of non-native speech production. Lingua 121, 1344-1358.


