Durational correlates for differentiating consonant sequences in Russian

Lisa Davidson
Kevin Roon
New York University
lisa.davidson@nyu.edu, kdroon@nyu.edu

In Russian, the same consonant sequences are permitted in various phonotactic environments. The presence of a word boundary or reduced vowel can be phonologically contrastive, and both learners and experienced listeners may rely on fine acoustic cues to discriminate between phonotactic possibilities. In this study, durational characteristics of consonant sequences are examined to establish whether speakers use duration to distinguish (a) word-initial clusters (#C1C2), (b) consonant–schwa–consonant sequences (#C1əC2), and (c) sequences divided by a word boundary (C1#C2). Both monolingual native Russian speakers and bilingual Russian–English speakers produced several types of target sequences: stop+consonant, fricative+consonant, and nasal+consonant. Results show that C2 is significantly longer in C1#C2 than in other sequences. For #C1C2, when C1 is a stop, there is no significant difference in duration when compared with other sequence types, though C1s of other manners are significantly shorter. Differences in C1 burst duration for stops are consonant-specific, but a longer interconsonantal duration is a reliable cue to schwa presence in #C1əC2. These results are discussed with respect to their implications for gestural coordination, segmentation, and language learning.

1 Introduction

Though obstruent–obstruent consonant sequences are relatively uncommon in the world's languages, they have proven to be an interesting subject for the study of consonant coordination in speech production (e.g. Browman & Goldstein 1988, 1990, 1995; Zsiga 1994; Byrd 1996; Byrd & Tan 1996; Gafos 2002; Davidson 2006a). Previous research has provided glimpses into how obstruent-initial consonant sequences are produced in several languages, demonstrating that there are several possibilities for articulatory and acoustic implementation. For example, in English, it has been shown that obstruents are weakly released before obstruents in word-medial and word-final position only about half of the time, and that the likelihood of release is dependent on the place of articulation of the first consonant in the sequence (e.g. *cactus*, Henderson & Repp 1982). Georgian contains so-called word-initial 'harmonic clusters', which are sequences of obstruents that must match in their laryngeal specification. Chitoran (1998) demonstrated that the first consonant of a harmonic cluster is typically released and is similar in length to its word-medial counterpart, indicating that the harmonic clusters are a sequence of consonants rather than a single complex segment. In Tsou, an Austronesian language that allows both word-initial and word-medial obstruent+consonant sequences,

Journal of the International Phonetic Association (2008) 38/2
doi:10.1017/S0025100308003447
© International Phonetic Association
Printed in the United Kingdom
https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

examination of stop releases and measures of fricative energy suggest that both types of sequences are released (stops) or less overlapped (fricatives) in initial position than in medial position (Wright 1996). Finally, in a study comparing the production of Russian and English consonant sequences across word boundaries, Zsiga (2003) found that when speaking their native languages, Russian speakers release stops in stop–stop sequences at word boundaries about five times as often as English speakers do.

This paper extends the examination of Russian consonant sequences by comparing consonant clusters across word boundaries with those in word-initial position, and with obstruent–schwa–consonant sequences. Russian is an interesting case for studying durational and other acoustic properties of obstruent-initial consonant sequences because there are fewer restrictions on their distribution than in many other languages. For example, consonant sequences like /tk/ or /pn/ can be found in word-initial position, word-medial position, and across a word boundary, as shown in (1). These consonants can also be found in consonant–schwa–consonant sequences, which differ from initial consonant clusters only by the presence of a short, reduced vowel.

(1)	Position	/tk/		/pn/	
	Initial	[tkɐnʲja]	'weaving'	[pnu]	'spurn'
	Medial	[vnɐklatku]	'with sugar'	[səvɐkupnə]	'jointly'
	Across word boundary	[nɐlʲit kɐnʲjak]	'the cognac is poured'	[pop nudnɨj]	'the priest is boring'
	CəC	[təkɐva]	'such'	[pənukatʲ]	'drive on'

Understanding how consonant sequences are implemented across various boundaries is important because it bears on several tasks that are required of speakers and hearers. First, it has been shown that the acoustic cues of ambiguous strings can be used by listeners to perform lexical segmentation. In some languages, like English, speakers prefer to use phonotactic cues to uncover the appropriate segmentation and only use acoustic cues as a secondary piece of information (Weber 2000). However, when a particular string of phonemes is ambiguous between two possible segmentations, listeners rely on acoustic cues like duration to distinguish the possibilities (e.g. *diep in* 'deep in' vs. *die pin* 'that pin' in Dutch or *night rate* vs. *nitrate* in English; Quené 1992, Hohne & Jusczyk 1994). Speakers also produce, and listeners use, durational cues to determine phrase boundaries (e.g. Lehiste 1960, 1972; Oller 1973; Klatt 1976; Scott 1982; Fougeron & Keating 1997; Cho & Keating 2001), and durational cues help listeners distinguish between a word that is an isolated form vs. being part of a polysyllabic word (e.g. *ham* vs. *hamster*; Salverda et al. 2003). Because there are few positional restrictions on consonant sequences in a language like Russian, it is possible that listeners are even more likely to rely on durational differences to perform lexical segmentation since phonotactic cues are less strong indicators of word boundaries.

Similarly, both second language learners, especially at early stages, and infants acquiring a first language may rely even more heavily on acoustic cues like duration for lexical segmentation than adult native listeners do even in non-ambiguous cases. While adult native speakers of a language rely on lexical/semantic information over both segmental/acoustic and metrical cues when all three types of information are available to them (Mattys, White & Melhorn 2005), early second language learners are less competent than native speakers at using top–down, contextual information in word recognition (Mayo, Florentine & Buss 1997). It is also possible that the lexicons of second-language learners are not developed enough to give rise to the process of lexical competition that has also been argued to aid segmentation (e.g. Norris 1994, Gow & Gordon 1995, Norris et al. 1997, McQueen 1998, Luce & Lyons 1999). Likewise, first-language learners have also been shown to rely on both phonotactics and on the acoustic information that mediates allophonic differences in order to perform word recognition tasks (Friederici & Wessels 1993, Hohne & Jusczyk 1994, Jusczyk, Hohne & Bauman 1999).

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

A second role that cues like duration may play is to assist both first and second language learners in uncovering the articulatory timing patterns of their target language. As has been pointed out by several studies, one aspect of second language learning is mastery of the articulatory coordination of sequential sounds in the language being acquired (Solé 1997, Tajima, Port & Dalby 1997, Zsiga 2003, Yanagawa 2007). In the production of sequential consonants, the articulatory coordination that is implemented affects whether the sequence is produced with close transition (the release of the first consonant is entirely overlapped by the second consonant), a release burst, or even a transitional schwa (Gafos 2002, Davidson 2005, Hall 2006, Bradley 2007). These audible cues to timing can give second language learners a clue as to the appropriate coordination between consonant sequences in their target language. As Zsiga (2003) pointed out, for example, proficient English learners of Russian are able to produce the Russian pattern of releasing pre-consonantal obstruents with relative ease whereas Russian speakers have considerably more difficultly with the close transition prevalent in English.

This paper focuses primarily on the role of duration in the production of three different kind of obstruent-initial sequences in Russian: word-initial consonant clusters (#CC), consonant sequences across a word boundary (C#C), and consonant–schwa–consonant sequences (#CəC). The first two types of sequences are the primary focus of interest, since for many consonant combinations, phonotactics alone would not be an adequate cue to lexical segmentation. If there are duration differences between the consonants in #CC vs. C#C sequences, listeners may use that information for segmentation and learners may deduce differences in coordination from the acoustic variation. The #CəC tokens are included as a comparison with #CC tokens, since a secondary question of interest is whether native speakers of Russian ever produce #CC words with a transitional, or 'excrescent', schwa between the consonants (see Davidson 2006a for production of these types of words by English speakers). While the production of excrescent schwas has never been examined in Russian, they have been noticed for other languages that have obstruent-initial consonant clusters, such as Piro (Matteson & Pike 1958), Moroccan Arabic (Gafos 2002), and Sierra Popoluca (Elson 1956). If second language learners are faced with variation in the input, it may impact how meticulously they acquire the coordination of non-native, obstruent-initial #CC sequences. While the purpose of this paper is not to determine how Russian listeners and second language learners of Russian use acoustic information to either segment speech or acquire coordination, the production data presented here are a necessary precursor for understanding what information in the signal is available to listeners.

A few differences between the consonant sequences under examination can be hypothesized on the basis of previous research. Two phenomena are expected to affect consonant length: domain strengthening, in which consonants at edges of domains are longer than non-edge counterparts (Fougeron & Keating 1997, Byrd 2000, Cho & Keating 2001), and the fact that consonants in clusters are shorter than singletons (e.g. Lindblom, Lyberg & Holmgen 1981). Thus, for example, both of the consonants in #CC are expected to be shorter than their counterparts in #CəC and C#C. For C2 specifically, there should be a decrease in duration from C#C, where C2 is at a word-boundary, to #CəC where it is a singleton, to #CC, where it is in a cluster. It is also possible that manner plays a role in consonant duration differences. If, as shown by Zsiga (2003), stops must always be released regardless of the type of consonant cluster and cannot be overlapped by the following consonant, then it might be expected that the duration of stops in C1 position in C#C and #CC should not differ.

A secondary comparison examined in this study is whether monolingual Russian and bilingual Russian–English speakers employ the same durational cues to distinguish between different types of consonant sequences. This question arose as a result of originally recruiting Russian speakers currently residing in the United States. Because these speakers are generally proficient bilinguals, it is possible that aspects of consonant sequence production in English are affecting their implementation of Russian consonant sequences (e.g. Flege & Hillenbrand 1984, Sancier & Fowler 1997). For example, since English medial and final obstruent clusters

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

are often produced without a release after the first consonant and final stops are also often unreleased (Henderson & Repp 1982, Zsiga 2003), it is possible that this coordination is visible in a bilingual Russian–English speaker's production in either #CC or C#C. This issue is ultimately important for language acquisition, since English-speaking students studying in the United States often have bilingual instructors providing a substantial amount of their input.

In the next section, the phonotactic properties of Russian consonant sequences relevant to this study are briefly presented.

1.1 Russian phonotactics

Russian is well known for having word-initial consonant sequences that violate traditional accounts of sonority well-formedness. For example, Russian contains words with both sonority plateaux and sonority reversals. Russian also allows several types of stop–fricative and obstruent–nasal sequences that are not common cross-linguistically (though they are found in other Slavic languages). These sequences can occur both with palatalized and non-palatalized consonants; the focus of this study is on sequences of non-palatalized consonants. All of these types of sequences are found in monomorphemic words, but word-initial obstruent–obstruent or obstruent–nasal sequences can also be formed through the concatenation of the prepositions /s/ 'with', /k/ 'to, toward', and /v/ 'in'. Examples of both monomorphemic and polymorphemic words are shown in (2).

(2)

a. Sonority plateaux		b. Sonority reversals		c. Stop–fricative		d. Obstruent–nasal	
[kto]	'who'	[zdɐroy̯]	'healthy'	[pʃɨno]	'millet'	[knʲigə]	'book'
[tkɐnʲja]	'weaving'	[ftɐroj]	'second'	[psɐlom]	'psalm'	[dno]	'bottom'
[fʃɨ]	'lice'	[lgatʲ]	'to lie'	[dvor]	'courtyard'	[gnɐitsa]	'to suppurate'
[k-tɐksʲi]	'toward the taxi'	[z-dɐvatʲ]	'to hand in'	[g-zɐrplatʲi]	'to the salary'	[k-mɐndatu]	'to the mandate'
[f-sɐfkʲe]	'in the trowel'	[f-tɐʃʲitʲ]	'to drag in'	[k-sɐbakʲi]	'towards the dog'	[k-nʲimu]	'towards him'

Two widely-described aspects of Russian phonology – namely, word-final obstruent devoicing and voicing assimilation in consonant clusters – must be mentioned in the context of the current study (Jones & Ward 1969, Halle 1971, Hamilton 1980, Avanesov 1984, Ozhegov 1990, Timberlake 2004). First, word-final obstruents (including /v/) are typically said to be devoiced in Russian, e.g. /porog/ → [pɐro̥g] 'threshold', /zdorov/ → [zdɐroy̥] 'healthy'. However, there is some evidence from preceding vowel duration, consonant duration, and duration of voicing into the closure that devoicing is not completely neutralized to a voiceless consonant (Pye 1986, Dmitrieva 2005).[1] (To indicate the possible ambiguity in the surface realization of word-final underlyingly voiced obstruents, they will be indicated in transcriptions by using the devoicing symbol, e.g. [y̥]). Second, obstruent clusters agree in voicing with the rightmost contiguous oral obstruent in the cluster, although an amplitude analysis of the clusters /st/, /sd/, /zt/ and /zd/ indicated the voicing of C1 was not always completely neutralized on the surface (Burton & Robblee 1997). While /v/ undergoes voicing assimilation, /v-katʲitʲ/ → [f-kɐtʲitʲ] 'to roll in', it does not trigger it, e.g. /svoboda/ → [svɐbodə] 'freedom'. Likewise, nasals do not trigger voicing assimilation, e.g. /knuta/ → [knuta] 'whip'. It is generally assumed that voicing assimilation does not cross

[1] We represent the assumed underlying representation in slashes, and the phonetic form in square brackets. Clitic boundaries are marked with a dash.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

word boundaries, though Timberlake (2004: 70) notes that it may be 'occasional between independent words, at least in close syntagms (and with connotations of colloquial register)'. Also, the exact nature and status of /v/ in Russian has been a subject of many studies, sometimes being analyzed as a fricative, sometimes as a sonorant. Because we do not take a stand on the exact nature of /v/, it is always examined as an independent category in the present study (Hayes 1984, Lulich 2004, Padgett to appear).

2 Experiment

2.1 Participants

The two types of participants included monolingual Russian speakers and bilingual Russian–English speakers. The six bilingual Russian speakers were primarily recruited from Russian communities in New York City such as Sheepshead Bay and Brighton Beach in Brooklyn. There were three male speakers and three female speakers. All of these speakers were born in the Moscow area and did not start learning English until they were between 12–17 years of age. These participants range in age from 20–29 years, with the exception of one 41-year-old female (who began learning English at age 16). All of the bilingual speakers report speaking both English and Russian, though they range from speaking English 100% of the time in a day to 60% Russian, 40% English. All but one report that, overall, they speak in English more than in Russian. It should be noted that while these speakers are L2 learners of English, we use the term 'bilingual' to indicate their proficiency in English, their frequent use of the language, and to simplify the terminology. The bilingual speakers were compensated for their time.

The five monolingual speakers were all recruited and recorded in Moscow. They range in age from 31–59 years. There were three female and two male speakers. These speakers reported learning either English or French, or German in high school, but none of them currently study or speak any of these languages. The monolingual speakers did not receive monetary compensation.

2.2 Materials

The reading materials consisted of phrases containing words with 16 different consonant cluster types (C1+C2). For each cluster type, there are three different types of sequeences: word-initial consonant cluster (#CC), consonant sequences across a word boundary (C#C), and consonant–schwa–consonant sequences (#CəC). The cluster types fall into 11 different combinations overall: /v/+fricative (two combinations), /v/+nasal (two combinations), /v/+stop (one combination), fricative+/v/ (two combinations), fricative+fricative (two combinations), fricative+nasal (two com-binations), fricative+stop (five combinations), stop+/v/ (three combinations), stop+fricative (four combinations), stop+nasal (five combinations), and stop+stop (five combinations). The consonants which could occur in C1 position were the fricatives /v f s z ʃ/ and stops /p t k d g/. C2 consonants were the fricatives /v f s z ʃ ʒ/, stops /p t k b d g/, nasals /m n/.

In all cases, both C1 and C2 were always non-palatalized consonants. To the extent that voicing assimilation within a cluster and across preposition boundaries is realized on the surface (cf. Burton & Robblee 1997), consonants in oral obstruent–obstruent sequences were either both voiced or both unvoiced on the surface, except where C1 was voiceless and C2 was /v/ or a nasal. In the C1+/v/ and C1+nasal cases, the initial obstruent could be either voiced or voiceless on the surface. Although a systematic analysis of word-final devoicing was beyond the scope of this study, impressionistically, underlyingly voiced C1s in the C#C condition were usually voiceless, although there were sporadic and infrequent instances of

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

voicing assimilation (see section 3.5 for further discussion). Some examples of stimuli are given in (3); the whole list of stimuli is found in appendix A.

(3)

	vd		**zn**	
#CC	/krʲitʃatʲ v-dogon/ [krʲiˈtʃatʲ vdɐˈgon]	‘to yell in pursuit’	/vrag znakomɨj/ [ˈvraɣ znɐˈkomɨj]	‘familiar foe’
C#C	/ivanov dovʲerʲil/ [ivɐˈnoʏ̯ dɐˈvʲerʲil]	‘Ivanov trusted’	/pʲrivʲoz navʲernoje vʲino/ [pʲrʲiˈvʲoz̥ nɐˈvʲernə vʲiˈno]	‘he probably brought wine’
#CəC	/zʲatʲ vod-o-laz/ [ˈzʲatʲ vədɐˈlaz̥]	‘the son-in-law is a scuba-diver’	/dosʲtʲig zanavʲeski/ [dɐˈsʲtʲig̥ zənɐˈvʲeski]	‘he reached/ attained the curtain’
	pʃ		**gb**	
#CC	/varʲitʲ pʃɨno/ [vɐˈrʲitʲ pʃɨˈno]	‘to cook millet’	/lʲestʲ k-baronu/ [ˈlʲestʲ gbɐˈronu]	‘to flatter the baron’
C#C	/utop ʃɨnok/ [uˈtop ʃɨˈnok]	‘the tavern was flooded’	/dʲenʲeg baronu/ [ˈdʲenʲig̥ bɐˈronu]	‘money for the baron’
#CəC	/tʃutʲ poʃɨvʲelʲitʲ/ [ˈtʃutʲ pəʃɨvʲiˈlʲitʲ]	‘to stir a little’	/otsenitʲ gabarit/ [ɐtsenˈitʲ gəbɐˈrit]	‘to evaluate the size’

In order to ensure that the speakers produced a schwa in the #CəC condition, it was necessary to carefully control the underlying vowel and the stress patterns in the phrases. In Russian, the vowels /o/ and /a/ are reduced to schwa only when they are two syllables before the stressed position, or after a stressed vowel. This is referred to as the ‘unstressed’ position. When /o/ and /a/ are immediately preceding a stressed syllable they also reduce, in this position to the vowel /ɐ/ (following Padgett & Tabain 2005). This is the ‘pre-stressed’ position. In word initial position, the first syllable of the #CəC words was always in unstressed position, followed by a pre-stressed /ɐ/. In order to keep the following vowel and the stress position consistent across sequence types, the #CC was always followed by a pre-stressed vowel. For the C#C words, we attempted to follow C2 with a pre-stressed vowel, but for three of the sequences (pn, dn, kn) it was impossible to construct meaningful, appropriate phrases, so in these cases C2 is followed by a stressed vowel. The statistical tests are modified to exclude these sequences when necessary. The relationship between the stress patterns of the three sequence types is summarized in (4).

(4)

	End of carrier	Unstressed	Pre-stressed	Stressed	Example
#CC	. . .C_x#	—	#C1C2ɐ	C'V. . .	[. . .fabrik tkɐnʲˈja]
#CəC	. . .C_x#	#C1ə	C2ɐ	C'V. . .	[. . .kak təkɐˈva]
C#C	. . .VC1#	—	#C2ɐ	C'V. . .	[. . .nɐˈlʲit kɐˈnʲjak]

Note: C_x indicates that the consonant preceding the target sequence at the end of the carrier was always the same.

The total number of phrases read by each participant was different for the bilingual and monolingual groups. The bilinguals were presented with 28 unique C1C2 combinations, each consisting of the three different sequence types, for a total of 84 target sequences. After beginning the analysis of the bilingual data, we noticed that it was occasionally difficult to find the boundary between the final (often palatalized) consonant of the carrier phrase and the beginning of C1 when C1 was a fricative. In these cases, we were concerned that we might be measuring the C1 fricative as being longer than it actually was, so for the monolinguals, we added another phrase for the #CC and #CəC condition where the carrier ended in a vowel

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

(note that for C#C, C1 was already always preceded by a vowel). Potential differences between the post-consonantal and post-vocalic fricative C1 are examined in the results section.

In addition to adding stimuli consisting of vowel-final carrier phrases for fricative-initial sequences, we also added five more C1C2 combinations for the monolinguals to record, for a total of 33 combinations. Thus, the monolinguals read 139 target sequences. The differences between the bilingual and monolingual stimuli are marked in appendix A.

2.3 Design and procedure

The target sequences were pseudo-randomized to ensure that no two phrases from the same C1C2 combination were next to one another. The phrases were then printed out in Cyrillic orthography in two columns on two sheets of paper. There were also two filler phrases at the beginning and end of each printed column to prevent the target phrases from being read with list intonation.

For both the monolinguals and bilinguals, at least one Russian-speaking experimenter was always in the room for the recording session. Before the recording began, participants were asked to read the list in front of the experimenter to ensure that they were familiar with all of the words and that the words were being read with the proper stress placement. The speakers were corrected if they did not place the stress on the right syllable. The experimenter remained with the speaker during the recording and if the speaker continued to make an error, she was prompted to repeat the phrase with the correct stress placement.

The bilingual speakers were recorded in a sound-proof booth at New York University. They held the sheets of paper containing the stimuli and read into a Shure Beta 85A microphone attached to a Marantz PMD-670 digital recorder. The microphone was placed 2–3 inches from each speaker's mouth. The monolingual speakers were recorded in a quiet room in a private home in Moscow using the M-Audio Microtrack 24/96 and the stereo electret microphone that comes with the Microtrack. The Microtrack was also placed approximately 2–3 inches from each speaker's mouth. The speakers read the phrase list through two times.

2.4 Analysis

Once the phrases were recorded, the waveform and spectrogram of both repetitions were examined in Praat 4.406 (Boersma & Weenink 2006) to make several duration measurements: C1 closure, burst and aspiration of C1 together (if present), voiced vowel (if present), and C2 closure.

A few points about the measurement criteria should be mentioned. First, if C1 was a stop, the closure was measured from the onset of silence to the point at which aperiodic energy for the burst appeared. If C1 was a fricative, the measurement began at the onset of high frequency aperiodic energy, even if that energy began before the end of the preceding vowel or sonorant consonant. When a fricative in C1 position was preceded by a palatalized stop (only /f/ and /v/ were ever preceded by palatalized stops), the palatalization had much more intense energy than the C1 fricatives did and was usually clearly demarcated. The measurement for these /f/ and /v/ C1s started at the offset of palatalization. When a plain stop preceded a fricative C1, measurement of C1 began at the offset of the preceding burst which was higher intensity or had a more diffuse range of energy than the fricative C1. C1 fricatives preceded by both kinds of stops and a C#C stop–stop sequence are illustrated in appendix B.

The start of C2 was measured from the offset of the last glottal pulse of the schwa in #CəC sequences. In #CC and C#C sequences, when C2 was preceded by a stop, it was measured from the offset of the burst just as a C1 preceded by a stop was measured. For fricative C2s, measurements depended on the sequence. When a low-energy fricative was followed by a higher energy fricative (e.g. /fs/, /vz/), the higher energy was generally characterized by an abrupt, clearly demarcated onset. When C2 was /v/, in most cases it appeared as an approximant, so the beginning of C2 was measured at the onset of formant structure. Two

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

examples of fricative C2 sequences are shown in appendix B figures B1 and B2. In all cases, the offset of C2 was at the beginning of formant structure in the following vowel. This included fricative C2s, which were occasionally followed by a short period of near silence before the formants of the following vowel started. This silence was measured as part of the duration of the fricative. The second formant (F2) was chosen as the most reliable indicator of formant structure, since the first formant was sometimes confusable with the fundamental frequency. In a small number of cases, it was impossible to definitively determine the boundary between the two consonants, so these tokens were discarded.

The vowel in #CəC sequences was measured from the onset of formant structure after the release burst or frication to the offset of the formants. Here, too, F2 was the most reliable indicator of the onset and offset of formant structure. When the schwa was followed by a nasal, the vowel duration was measured until the abrupt decrease in energy around the second formant which is characteristic of a nasal. Two examples of #CəC sequences are shown in appendix B figures B4 and B5.

If vocalic material was present between an intended #CC or C#C sequence, the vowel duration was measured and coded as a vowel. The acoustic record was considered to have a vowel in this position only if formant structure was clearly present. The presence of a voice bar alone was not sufficient for coding as a vowel.

If the speaker paused between the two consonants in C#C tokens, or before a C1 stop in #CC or #CəC stimuli, they were not used in the statistical analysis.

3 Results

In this section, the results for both duration measurements and for information about the presence of vowels and the absence of bursts in #CC and C#C stimuli are presented. Prior to any further analyses, we compared the duration of C1 fricatives for the monolinguals to determine whether the stimuli in which C1 was preceded by a vowel were significantly different from the C1 fricatives that were preceded by a stop (see explanation in section 2.2). A univariate analysis of variance (ANOVA) with subjects as a random factor and preceding segment (stop or vowel) and fricative voicing (surface voiced or voiceless) as an independent variable showed no significant main effect for preceding segment [$F(1, 4) < 1$], a significant main effect of C1 voicing [$F(1, 4) = 44.31$, $p < .002$], and no interaction [$F(1, 4) = 3.06$, $p = .15$]. The significant effect of C1 voicing demonstrates that overall, voiced fricatives are shorter than voiceless ones, but since there was no significant main effect of preceding segment or interaction, preceding segment is collapsed in the remaining analyses.

All of the following analyses were carried out in SPSS. Three further points should be made. First, for each dependent variable analyzed, we treated monolingual vs. bilingual as an independent variable to determine whether there were any significant differences in the cues for each sequence type (i.e. #CC, C#C, #CəC). We call this variable LINGUALISM. Second, for each analysis, we examined both the whole data set and the subset of the data in which the vowel following the target sequence is always pre-stressed (see section 2.2). Because the stress of the syllable containing the target sequence may affect durations of various constituents of the target sequence, we wanted to ensure that potential differences were not obscured by non-stress-matched items. When the stress-matched data has a significant or insignificant pattern that is not mirrored in the whole data set, these results are reported. Otherwise, it can be assumed that the acoustic measures were the same across the two sets of the data. Third, in all of the following analyses that include consonant manner as an independent variable, we treated /v/ as a category separate from all other fricatives. This is because /v/ in Russian can be realized either in a more fricative-like manner or as an approximant, and changes in duration may be realized differently than other fricatives.

Because the number of observations in each cell is not always the same, we used Type III Sums of Square to correct for any discrepancies.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

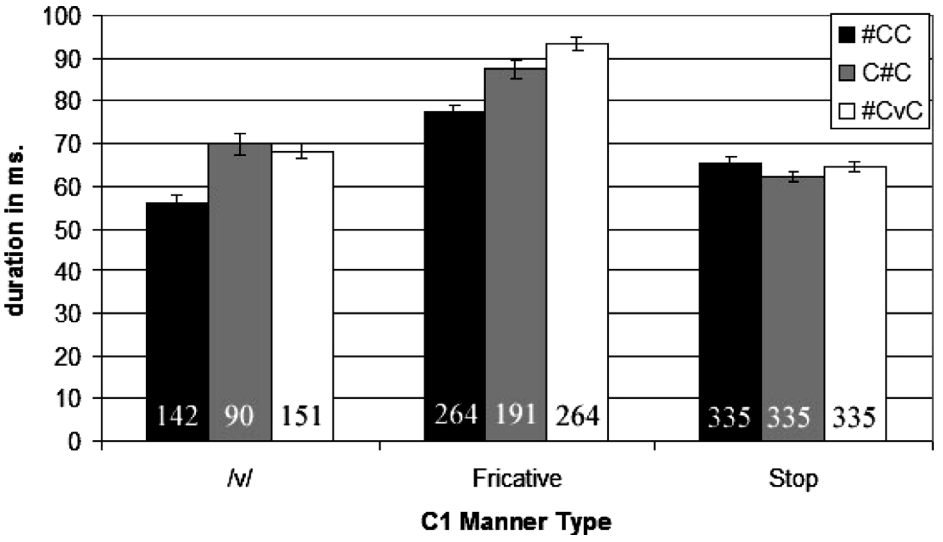


Figure 1 Duration in milliseconds of C1 for each manner and sequence type. #CvC = #CəC. Error bars represent standard error. In this and future figures, the number in each bar represents the number of tokens for each manner and sequence combination.

3.1 Total sequence duration

An ANOVA was carried out to determine whether the mean durations of the whole target sequences are significantly different. Independent variables were sequence type (#CC, C#C, #CəC) and lingualism (bilingual, monolingual). Results indicated that there was a significant main effect of sequence type [$F(2, 1982) = 106.08$, $p < .001$], and lingualism [$F(1, 1982) = 46.94$, $p < .001$]. The interaction was not significant. Scheffé post-hoc tests show that as expected, #CəC (N = 709, mean = 224 ms) is significantly longer than either #CC (N = 702, mean = 164 ms) or C#C (N = 589, mean = 180 ms) ($p < .001$), and that #CC and C#C are significantly different from one another ($p < .003$). There is a main effect of lingualism because the mean durations collapsing across all sequences for bilinguals is 177 ms (N = 982), whereas the mean duration for monolinguals is 202 ms (N = 1018).

3.2 C1 duration

An ANOVA examining differences in C1 duration was carried out. Independent variables were sequence type, manner of C1 (stop, fricative, /v/), and lingualism. Results showed main effects of all three independent variables [sequence type: $F(2, 2095) = 25.85$, $p < .001$; C1 manner: $F(2, 2095) = 171.85$, $p < .001$; lingualism: $F(1, 2095) = 14.68$, $p < .001$]. Significant two-way interactions were manner by lingualism [$F(2, 2095) = 5.94$, $p < .003$] and sequence by manner [$F(4, 2095) = 12.26$, $p < .001$]. The three-way interaction was not significant.

The ANOVA for the stress-matched subset showed almost all of the same results, except that the interaction between manner and lingualism was only marginal [$F(1, 1713) = 2.47$, $p = .08$]. (Ns: #CC = 624, C#C = 499, #CəC = 608; fricatives = 589, stops = 759, /v/ = 383.)

For the general data set, the graph in figure 1 shows that the significant main effect of C1 manner was due to duration differences between /v/, fricatives, and stops. A Scheffé post-hoc test indicated that collapsing over sequence type, fricatives were significantly longer than /v/ and stops ($p < .001$), which were not significantly different from one another. Collapsing over manner, the C1 of #CəC was the longest (mean = 75 ms), followed by C#C (mean = 73 ms) and #CC (mean = 65 ms). All comparisons are significant: #CC vs. #CəC ($p < .001$), #CC vs.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

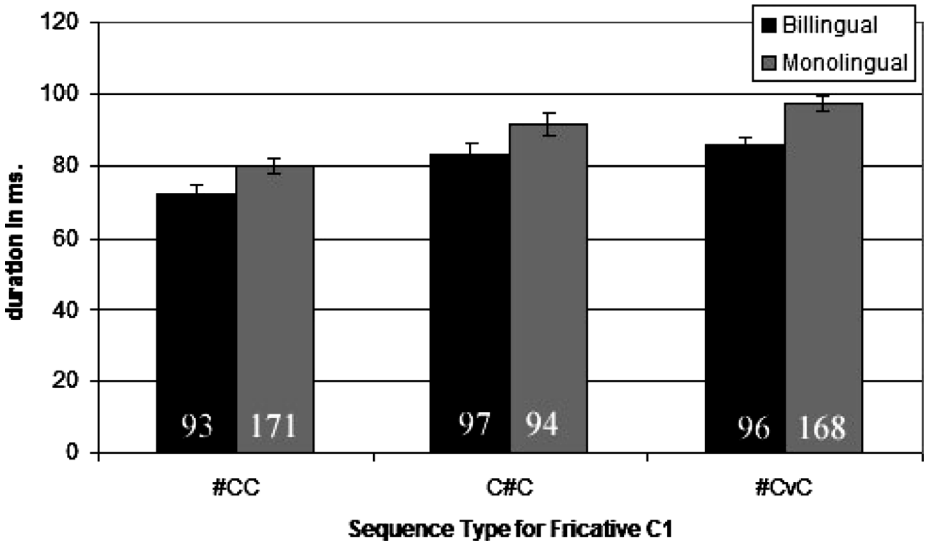


Figure 2 Comparison of bilinguals and monolinguals on the duration of fricative C1s by sequence type.

C#C ($p < .042$), #CəC vs. C#C ($p < .005$). Nevertheless, little importance can be ascribed to the significant difference between #CəC and C#C, since 2 ms is below the threshold of the just noticeable difference (JND) for consonant duration, which is approximately 10 ms, but depends on particular manners (Klatt 1976).

Separate univariate ANOVAs were carried out for each C1 manner, collapsing over lingualism. For both /v/ and fricative C1s, there was a significant effect of sequence type [/v/: $F(2, 377) = 17.38$, $p < .001$; fricative: $F(2, 713) = 21.83$, $p < .001$]. Scheffé post-hoc tests showed that for both of these manners, C1 was significantly shorter in #CC sequences than C#C or #CəC ($p < .001$), which were not significantly different from each other. When C1 was a stop, there were no differences in duration based on sequence type [$F(2, 1005) = 1.71$, $p = .18$].

A separate ANOVA for each of the C1 manners was also carried out with lingualism and sequence type as independent variables. Results show that the only manner for which the bilinguals differed from the monolinguals was for the fricatives [$F(1, 713) = 19.95$, $p < .001$]. The main effect of sequence was also significant [$F(2, 713) = 21.83$, $p < .001$], but the interaction was not. The comparison across fricative C1s was also significant for the stress-matched data [$F(1, 583) = 10.81$, $p < .001$]; no significant differences were found for the stops or /v/ in the stress-matched data. As shown in figure 2, the monolinguals' fricatives were consistently longer than the bilinguals' fricatives.

In sum, the results indicate that C1 duration was significantly longer for C#C and #CəC than for #CC for /v/ and fricatives, and there was no difference for stops. For monolinguals, the duration of fricatives was longer overall, but since there was no interaction with sequence type, the pattern is the same as the general finding.

3.3 C2 duration

Duration of C2 was examined with an ANOVA including sequence type, manner of C2 (stop, fricative, /v/, nasal), and lingualism as independent variables. Results showed main effects of all three independent variables [sequence type: $F(2, 2099) = 28.38$, $p < .001$; C2 manner: $F(3, 2099) = 357.92$, $p < .001$; lingualism: $F(1, 2099) = 19.04$, $p < .001$]. None of the two-way

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

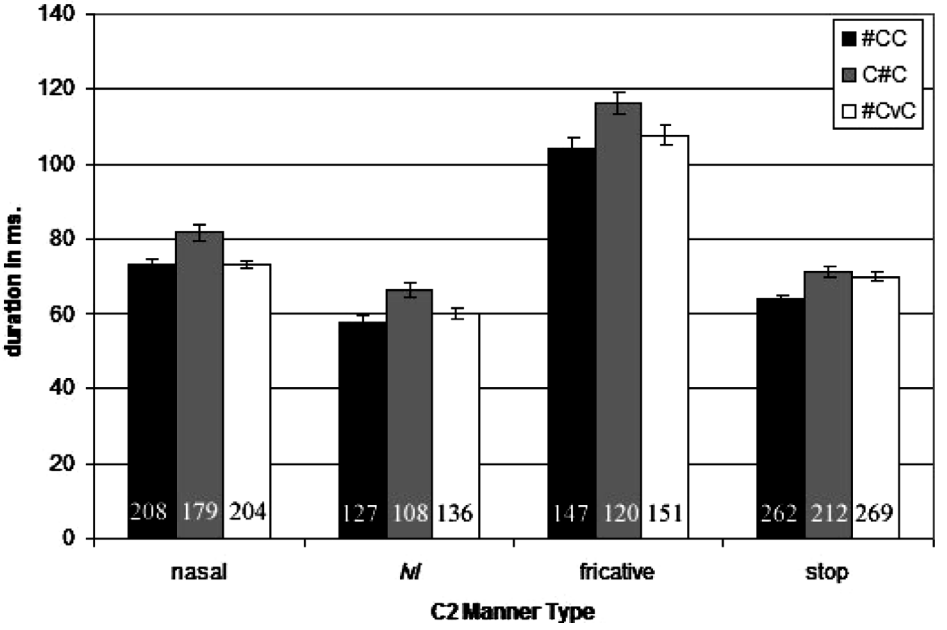


Figure 3 Duration of C2 for each manner and sequence type.

interactions were significant, although manner by lingualism was marginal [$F(3, 2099) = 2.30$, $p = .075$]. The three-way interaction was marginal [$F(6, 2099) = 1.89$, $p = .078$].

The results for the stress-matched subset were the same, except that the three-way interaction was not even marginal. (Ns: #CC = 627, C#C = 502, #CəC = 609; nasals = 404, fricatives = 353, stops = 631, /v/ = 350.)

The graph in figure 3 shows that the significant main effect of C2 manner for the whole dataset was due to duration differences between all four C2 manner types. A Scheffé post-hoc test indicates that collapsing over sequence, the fricatives were the longest consonants (mean = 109 ms) followed by nasals (mean = 76 ms), stops (mean = 68 ms), and /v/ (mean = 61 ms). All comparisons were significant ($p < .001$). For sequence type, C2 of C#C was the longest (mean = 84 ms), followed by #CəC (mean = 77 ms) and #CC (mean = 74 ms). All comparisons were significant: #CC vs. #CəC ($p < .021$), #CC vs. C#C ($p < .001$), #CəC vs. C#C ($p < .001$), though again, the difference between #CəC and #CC may be below the JND threshold. Since the interaction of manner by sequence was not significant, these patterns hold up for all C2 manner types.

An ANOVA was run separately for each of the C2 manner types with lingualism and sequence as independent variables. For /v/, there were significant main effects for sequence [$F(2, 365) = 8.90$, $p < .001$] and lingualism [$F(1, 365) = 19.06$, $p < .001$], and a significant interaction [$F(2, 365) = 3.00$, $p < .05$]. As shown in the graph in figure 4, monolinguals produced consistently longer /v/ for all sequence types, although only the difference for #CC was significant ($p < .001$).

For fricatives, there was a main effect for sequence type [$F(2, 412) = 5.50$, $p < .004$] and a marginal effect for lingualism [$F(2, 365) = 3.00$, $p < .06$]. The main effect for lingualism in the stress-matched case was significant [$F(1, 347) = 4.86$, $p = .028$]. There was no two-way interaction for either data set. Again, figure 4 shows that fricatives for monolinguals were significantly longer.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

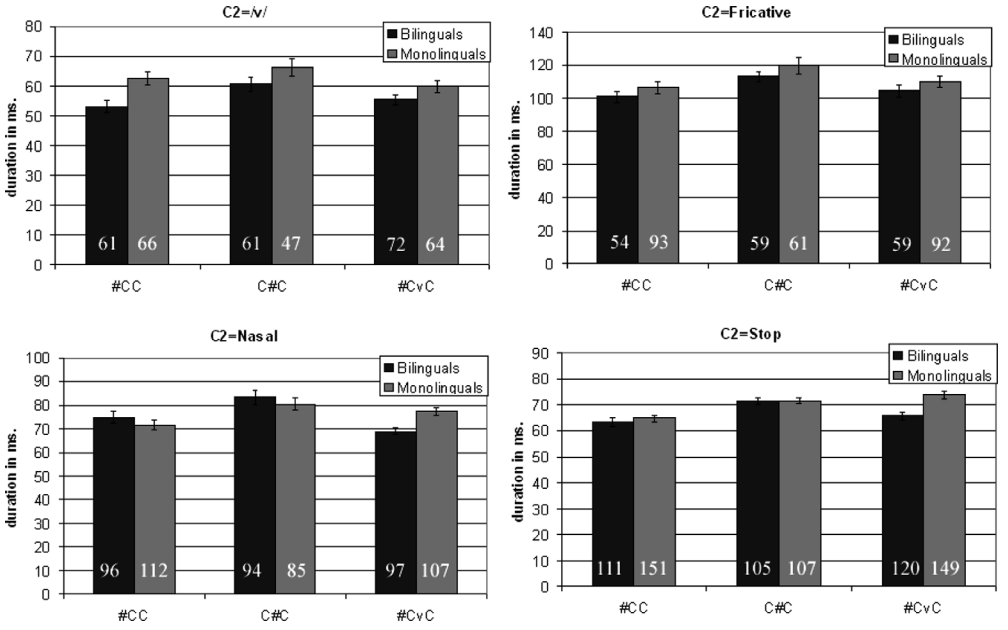


Figure 4 Comparison of bilinguals and monolinguals on the duration of each C2 by sequence type.

For nasals, there was a main effect for sequence type [$F(2, 585) = 8.66$, $p < .001$] but no main effect for lingualism [$F(1, 585) < 1$]. However, the interaction between sequence type and lingualism was significant [$F(2, 585) = 4.11$, $p = .017$]. Pairwise comparisons showed that this difference was only significant within #CəC sequences ($p < .001$), as illustrated in figure 4.

Finally, for stops, there were main effects for sequence type [$F(2, 737) = 10.81$, $p < .001$] and lingualism [$F(1, 737) = 5.36$, $p = .021$]. The interaction was also significant [$F(2, 737) = 3.48$, $p = .031$]. As indicated by figure 4, the only significant difference was for #CəC ($p < .001$) ($p < .05$ for stress-matched data), where monolinguals produce significantly longer C2 stop closures than bilinguals.

In sum, the results for C2 demonstrate that C2 duration was longest for C#C, followed by #CəC, which was still longer than #CC, though these differences should be treated with caution. Notably, C2 of C#C for all manners was reliably longer than C2 for #CC. The findings for lingualism demonstrate that monolinguals tended to have slightly longer C2 consonants than the bilinguals, which was reflected to varying degrees for all C2 manner types.

3.4 Vowels and bursts

Another measure of interest in this study was whether native Russian speakers ever inserted vocalic material between the consonants in either C#C or #CC sequences. A count of these 'excrescent schwa' productions is shown in table 1. The sequence with most excrescent schwas was /pn/ ($N = 10$), followed by /gn/ ($N = 8$), /dv/ ($N = 6$), /gb/ ($N = 5$), /gd/ ($N = 5$), and /vn/ ($N = 5$). All other sequences with excrescent schwas had three or fewer tokens, but most of them were composed of voiced oral obstruents (including /v/) or voiced obstruent+nasal. If excrescent vowels are a result of less overlapping of the consonant constriction than is usually implemented (Davidson 2006a), then it is not surprising that they occur most often between voiced consonants since it may be easier to simply continue vibrating the vocal folds in that environment.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

Table 1 Counts of excrescent schwas.

	#CC	C#C
Bilinguals	19 (out of 323)	2 (out of 322)
Monolinguals	40 (out of 379)	10 (out of 267)

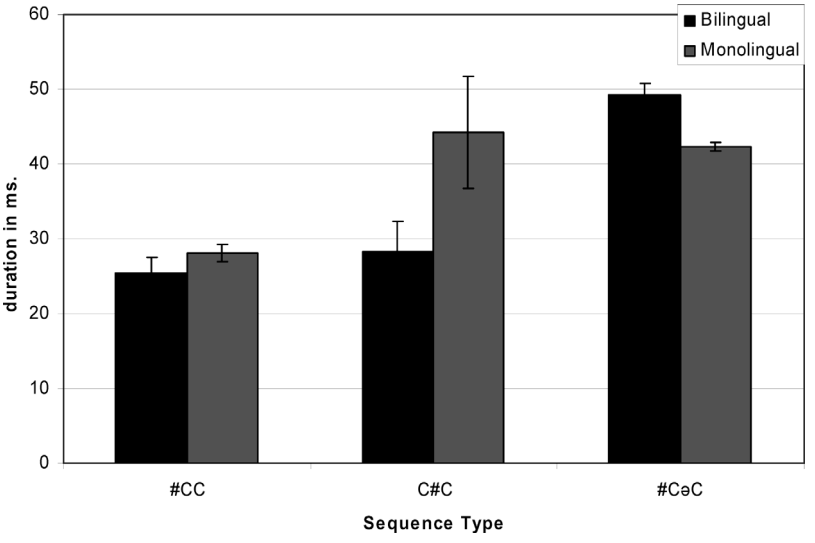


Figure 5 Comparison of the duration of excrescent (#CC, C#C) and lexical (#CəC) vowels for bilinguals versus monolinguals.

An ANOVA was carried out to determine whether the duration of the vowel was significantly different if it was a lexical vs. excrescent vowel. In order to minimize the unevenness in the comparison (since the #CəC sequence type contains far more vowels than the C#C and #CC types), the tokens in the #CəC condition were limited to those found to contain excrescent schwas in the C#C and #CC types. Independent variables were sequence type and lingualism. Results showed a significant effect of sequence [$F(2, 543) = 44.86$, $p < .001$] and no effect of lingualism [$F(1, 543) = 1.09$, $p = .30$]. The interaction was significant [$F(2, 543) = 4.89$, $p < .008$]. As illustrated in figure 5, post-hoc tests indicate that the significant main effect of sequence was due to a longer vowel being produced for #CəC tokens than for #CC ($p < .001$). Somewhat surprisingly, there was no difference between #CəC and C#C tokens, nor was the difference between bilinguals and monolinguals in the C#C condition significant. The lack of effect, and the long duration of the excrescent vowels for bilinguals in the C#C condition is most likely due to the very small sample size.

In addition to vowel duration, duration of the burst following stops in C1 position was also examined. In this case, because burst duration may be especially vulnerable to the stress of the following syllable, the analysis was restricted to stops in the subset of stress-matched stimuli. An ANOVA was carried out with sequence type, C1 type (/p t k d g/) and lingualism as independent variables. Results showed a marginal main effect of sequence type [$F(2, 672) = 2.93$, $p = .054$], a main effect of C1 type [$F(4, 672) = 29.11$, $p < .001$], but no effect of lingualism [$F(1, 672) < 1$]. The only significant interaction was the two-way interaction between sequence and C1 [$F(8, 672) = 6.44$, $p < .001$]. Scheffé post-hoc tests indicate that the main effect of sequence is due to a significant difference between #CC (mean = 23 ms) and C#C (mean = 26 ms), but this very small difference is below the

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

Table 2 C1 burst durations for each sequence and C1 type in milliseconds. Significant differences between sequence types according to Scheffé post-hoc tests are shown below the mean duration values (> = significantly larger).

	C1 type				
	d	g	p	k	t
#CC	19.6	20.5	10.2	28.8	26.4
#CəC	17.1	20.2	16.3	33.5	18.3
C#C	27.5	26.6	14.6	27.3	24.1
	C#C > #CC, #CəC	C#C > #CC, #CəC	#CəC > #CC	#CəC > #CC, C#C	#CC > #CəC

JND threshold (for comparison, mean C1 burst duration for #CəC = 24 ms; Ns: #CC = 223, C#C = 221, #CəC = 258; /d/ = 90, /g/ = 237, /k/ = 269, /p/ = 47, /t/ = 59).

Collapsing over lingualism, the two-way interaction between sequence and C1 is due to a complex series of differences among the consonants. These are shown in table 2. One consistent result is that the burst for the voiced stops are consistently longer in the C#C condition. This suggests that voiced stops are often devoiced in word-final position, resulting in the longer burst associated with voiceless stops (see section 3.5 for further discussion of the status of final devoicing in this data). For the voiceless stops, /p/ and /k/ pattern similarly, where the burst of C1 in #CəC is significantly longer than at least one of the consonant cluster types, but for /t/ the C1 burst #CC is longer than #CəC. There is no immediately apparent explanation for why /t/ is different from /p/ and /k/, but the lack of a difference between C#C and #CC for all of the voiceless stops is the most important outcome for this study, since it indicates that C1 burst duration does not distinguish between #CC and C#C when C1 is voiceless.

In addition, Russian speakers occasionally failed to produce a burst in C#C and #CC sequences. Overall, there were 50 instances of a missing burst for stop–initial #CC sequences (15%, 324 total) and 54 missing bursts for C#C sequences (17%, 322 total). This number is slightly lower than reported in Zsiga (2003), who found that 27% of non-identical stop#stop sequences had missing releases for the first consonant, but the current study included stop–stop, stop–fricative, and stop–nasal sequences. Thus, it appears that Russian speakers are fairly consistent in releasing stops before other consonants. Likewise, the vowel was present in all but 18 cases (out of 781, 2%), indicating that Russian speakers do not delete the vowel or overlap the surrounding consonants and the vowel such that it would not be present in the acoustic record. This is less than the English speakers examined by Davidson (2006b), who either deleted pre-tonic schwa or reduced it to aspiration at a normal speech rate in approximately 11% of the tokens.

3.5 Voiceless C1 subset

The analyses in the preceding section collapse over sequences that would be traditionally characterized having two voiced or two voiceless members, except in the C#C condition, where word-final devoicing usually causes a C1 that is underlyingly voiced to surface as voiceless. However, it is possible that the voiced C1 in C#C may not always be neutralized to a voiceless obstruent. First, word-final devoicing of obstruents may not be complete or occur in 100% of utterances, which could lead to an ambiguity between C#C and #CC for sequences of voiced consonants that is similar to that present for voiceless sequences. Second, this situation may be further compounded by regressive assimilation across word boundaries, which would also lead to an underlyingly voiced obstruent in C#C position surfacing as voiced. As noted in section 2.2, a voiced C1 in C#C tokens was impressionistically rare, but figure 6 is an example of the utterance /iz-'tʃʲez # bolʲ'noj/ where the word-final /z/ clearly surfaces as [z].

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

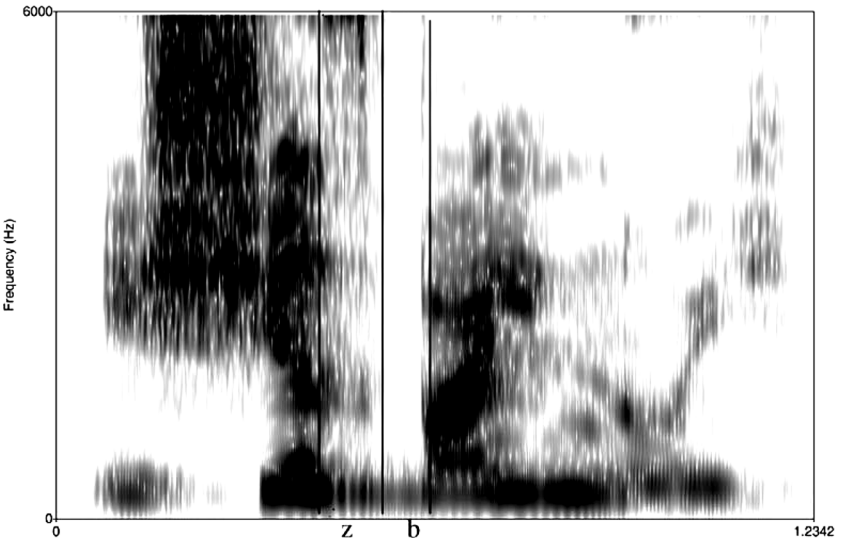


Figure 6 Spectrogram of a production of /iz-'tʃʲez # bolʲ'noj/ 'the sick man disappeared' by monolingual speaker 1110. The /z/ and /b/, identified on the spectrogram, are both voiced even though the /z/ is word-final.

In cases where word-final devoicing is realized, a listener who does encounter a voiceless–voiced obstruent sequence (except when C2 is /v/ or a nasal) is presented with a salient cue that they are hearing a word-boundary and not an initial consonant cluster. Thus, the following analyses examine only the subset of data that contains the 15 sequences with a voiceless C1 followed by a voiceless obstruent, /v/, or a nasal (e.g. /tk/, /kv/, /pn/, /ft/). It is an interesting question as to whether this subset follows the general pattern for the implementation of consonant sequences, or whether there are any differences that further distinguish C#C vs. #CC in these especially ambiguous cases. We examine this question by looking at C1 duration, and C2 duration. C1 burst duration is not analyzed again, since the breakdown by C1 type in section 3.4 provides the relevant information.

3.5.1 C1 duration

C1 duration was analyzed with an ANOVA with the independent variables sequence type, manner of C1 (stop, fricative), and lingualism. Results showed main effects of all three independent variables [sequence type: $F(2, 932) = 8.73$, $p < .001$; C1 manner: $F(1, 932) = 133.89$, $p < .001$; lingualism: $F(1, 932) = 17.06$, $p < .001$]. Significant two-way interactions were manner by lingualism [$F(1, 932) = 9.84$, $p < .002$] and sequence by manner [$F(2, 932) = 12.95$, $p < .001$]. The three-way interaction was not significant.

The results for the stress-matched data were the same except that manner by lingualism was not significant.

Scheffé post-hoc tests showed that #CəC (mean = 80ms) was significantly different from both #CC (mean = 73 ms) and C#C (mean = 70 ms) ($p < .002$), but that these were not different from one another. As illustrated in figure 7, for fricatives C1 of #CəC was significantly longer than C#C or #CC ($p < .05$ and $p < .001$, respectively), which were not significantly different from each other. For stops, C1 of C#C was significantly shorter than C1 of #CC ($p < .05$), but the duration of C1 of #CəC was not significantly different from either of the other sequence types. For stress-matched data, there were no differences of sequence type on C1 closure duration when C1 is a stop.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

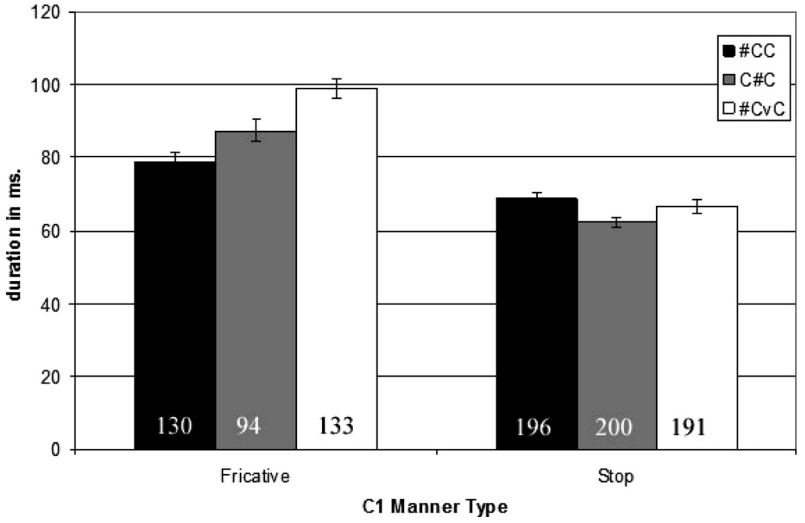


Figure 7 Duration of C1 for each manner and sequence type for the voiceless C1 subset of the data.

3.5.2 C2 duration

Duration of C2 was examined with an ANOVA with sequence type, manner of C2 (stop, fricative, /v/, nasal), and lingualism as independent variables. Results showed main effects of all three independent variables [sequence type: $F(2, 934) = 28.61$, $p < .001$; C2 manner: $F(3, 934) = 433.00$, $p < .001$; lingualism: $F(1, 934) = 20.01$, $p < .001$]. The interaction between sequence and manner was significant [$F(6, 934) = 5.73$, $p < .001$], as was the interaction between lingualism and manner [$F(3, 934) = 3.73$, $p < .01$]. The three-way interaction was not significant.

The results for the stress-matched subset were the same, except that the two-way interaction between lingualism and manner was not significant [$F(3, 675) = 1.71$, $p = .17$].

The significant main effect of sequence type is the same as the general data set: C2 of C#C is the longest (mean = 91 ms), followed by #CəC (mean = 83 ms) and #CC (mean = 78 ms). A Scheffé post-hoc test confirms that all comparisons were significant: #CC vs. C#C ($p < .001$), #CC vs. #CəC ($p < .012$), and #CəC vs. C#C ($p < .001$). Unlike the general data set, there was an interaction between sequence and manner. As shown in figure 8, the significant differences are as follows: for /v/, C2 of C#C was significantly longer than #CC or #CəC ($p < .05$ and $p < .005$, respectively), which were not significantly different from each other. When C2 was a fricative, C2 of C#C was significantly longer than C2 of #CC ($p < .005$). When C2 was a nasal, C#C was significantly longer than #CC or #CəC (both $p < .001$), which are not significantly different from each other. However, for the stress-matched data for nasals, there was no effect of sequence type ($p = .416$). When C2 was a stop, #CC was significantly shorter than C#C or #CəC ($p < .05$ and $p < .001$), respectively for all data, which were not significantly different from each other.

4 General discussion

The results of the acoustic analysis indicate that speakers produce a number of consistent durational differences demarcating C#C, #CC, and #CəC sequences in connected speech. The differences are similar regardless of whether the speakers are monolingual or bilingual.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

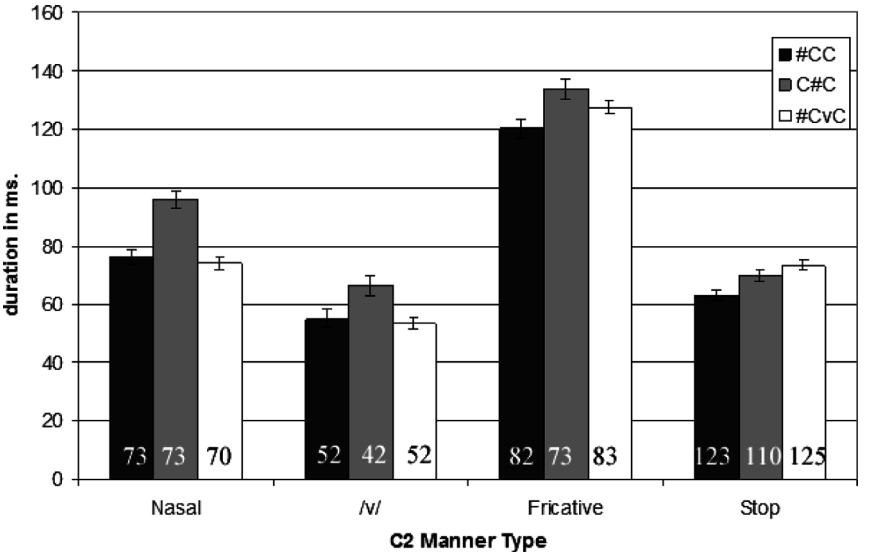


Figure 8 Duration of C2 for each manner and sequence type for the voiceless C1 subset of the data.

Likewise, there are no substantial differences between the whole dataset, the stress-matched subset and the voiceless C1 subset. This indicates that durational cues distinguishing C#C, #CC, and #CəC sequences are general ones, and are not enhanced by speakers for sequences that are ambiguous and more likely to be confusable to the listener.

For all speakers, unsurprisingly, #CəC sequences are the longest of all, and C#C sequences are longer than #CC sequences. The relative shortness of #CC sequences is largely attributable to C1 duration results, which show that for /v/ and fricatives, #CC is shorter than the other two types. Following Browman & Goldstein (1995), this decrease in C1 duration in #CC may be due to greater overlap in consonant clusters within the same syllable position.[2] However, the results also showed that there are no differences for stops. These results are largely mirrored in the voiceless C1 subset, where C1 of #CC is shortest for the fricatives, but is either equal in length or even a little longer than the other sequences when C1 in #CC is a stop. That stops in C1 position do not differ in length depending on sequence type is consistent with the canonical coordination in Russian which ensures that initial stops in consonant sequences are always released (Zsiga 2003). Because stops must be released before obstruents and nasals, increased overlap is prevented in those sequences.

The findings for C2 for all data sets indicate that the duration of the second consonant of C#C sequences is the longest, followed by #CəC and then by #CC. To the extent that the difference between #CəC and #CC is meaningful since it may be below the JND threshold, these results are easily explained. Notably, C2 of C#C sequences is word-initial, C2 of #CəC is syllable-initial but not word-initial, and C2 of #CC is in a consonant cluster. These results are consistent with reports of domain strengthening (Fougeron & Keating 1997, Byrd 2000,

[2] Note that these results contradict Byrd's (1996) findings for the sequence /sk/ in English. Her examination of consonant overlap using electropalatography showed that there was less overlap for /#sk/ than for /s#k/. There are a number of reasons for a differences: (1) English and Russian may simply have different coordination patterns, (2) Byrd looked only at /#sk/, whereas other fricative-initial sequences were examined in this study, and (3) /s/ may have a different coordination pattern than other consonants, which may be related to why /s/ is so common in onset clusters whereas other fricatives are not (Morelli 1999).

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

Cho & Keating 2001), and with the finding that consonants in clusters are shorter than singletons (e.g. Lindblom et al. 1981).

While the durations of C1 and C2 distinguish between different types of consonant sequences, burst durations are less consistent. In the stress-matched data, which control for any possible effects of stress placement on aspiration duration, the differences between the sequences for voiced stops indicate that the C1 burst of an underlyingly voiced stop in the C#C sequence is longer than the burst in #CC or #CəC, which is consistent with word-final devoicing. For the voiceless stops, there were no differences in burst length between C#C and #CC. As noted in section 3.5, if a listener encounters a voiceless+voiced obstruent sequence, however variable its realization may be, that particular utterance should be unambiguous to the listener as a word-boundary. But because word-final devoicing may not be complete and regressive voicing assimilation may cross word-boundaries (Burton & Robblee 1997, Timberlake 2004, Dmitrieva 2005), these unambiguous sequences may not be encountered as often as might be expected. Thus, C1 burst duration alone does not adequately distinguish between the consonant sequences examined in this study.

The comparison of monolinguals and bilinguals indicates that there are few consequential differences between the two types of speakers with regard to the relationship between different types of sequences. The main difference between monolinguals and bilinguals is that the monolinguals have longer overall fricative durations for both C1 and C2. For nasals and stops, the monolinguals also produce C2 of #CəC with a longer duration than bilinguals. One possible sources of these differences is that the inherent differences in consonant length for these sequences have been modified through contact with English (cf. Sancier & Fowler 1997). Regardless of the source of the differences between monolinguals and bilinguals, none of them affect the relationship between #CC and C#C, indicating that the cues to differentiate these types of consonant sequences are the same for both types of speakers.

These results, along with the phonological facts presented in the introduction, demonstrate that there are several phonetic properties of Russian that may be combined to assist the listener in segmenting the speech stream and the learner in identifying new lexical items. First, when word-final devoicing does happen, listeners faced with a voiceless–voiced obstruent sequence can be sure that it is occurring across a word boundary. However, there is a considerable number of sequences which can occur in either #CC or C#C position, so the consistent distinctions between #CC or C#C as reflected mostly in C1 and C2 duration differences should be available to the listener for perceptual tasks. The distinction between stop-initial sequences relies more heavily on C2 duration, since the requirement that C1 stops be released prevents the shortening of those consonants.

A secondary question of interest in this study was whether native Russian speakers ever produce #CC or C#C with an excrescent schwa. Since these two types of sequences are lexically contrastive with #CəC, we did not expect to find a significant number of excrescent schwas. When the relatively few instances of excrescence did occur, the schwas are shorter than lexical ones. Though no further acoustic analysis of the excrescent schwas was carried out (e.g. formant measurements), the shorter duration is consistent with previous experimental work indicating that these schwas result from decreased articulatory overlap of the consonants in the sequence rather than insertion of a vowel (Davidson 2005, 2006a). Russian speakers exhibit this overlap pattern infrequently, suggesting that they generally adhere to a constrained range of consonant coordination for obstruent–obstruent and obstruent–nasal sequences.

The issue of excrescent schwas is important when considering how second language learners may use the acoustic signal to determine the coordination of consonant sequences in Russian. For example, learners of Russian, who may initially either epenthesize a schwa between the two consonants in an initial cluster or produce them with insufficient overlap, must ultimately use cues from the signal to learn that Russian obstruent-initial #CC sequences are overlapped enough to exclude any vocalic portion between the two consonants. Since there are very few instances of excrescent schwas in the input, such schwas on the part of language learners would be attributable to their own phonological and phonetic systems, not to the input

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

from the Russian speakers. Likewise, non-native speakers must also use the cues in signal to learn that whether stops are in clusters (#CC) or in word-final position (C#C), the duration of the release burst is the same. As shown by Zsiga (2003), this is not particularly difficult for English speakers to learn for word-final stops, but it is likely to be a problem even for proficient English learners of Russian to produce most obstruent-initial #CC sequences with the right amount of overlap since these sequences are phonologically prohibited in English. In fact, previous research has shown that native English speakers have trouble producing even fricative-initial consonant sequences with close coordination in word-initial position (Davidson 2005, 2006a).

The findings of this study and the preceding discussion point the way to at least two areas for future study. First, now that acoustic differences between the consonant sequences under investigation have been established, the next step would be to determine how salient they are to both native language and second language learning listeners. If listeners use these acoustic cues in speech segmentation, then we would expect that they could distinguish between . . . V#CCV . . . and . . . VC#CV . . . sequences that are either lexical minimal pairs or that form nonce words. Likewise, it might be expected that there is a correlation between the perceptual abilities of second language learners who can distinguish between such pairs and the acoustic properties of their own implementation of this distinction (e.g. Bradlow et al. 1996).

Second, whereas Russian has a three-way distinction between #CC, C#C, and #CəC, most other Slavic languages do not have the latter category since schwa is neither a phoneme nor an allophone in their vowel inventories. Thus, while learners of Russian should ultimately learn that the insertion of an excrescent schwa in #CC sequences would acoustically compete with #CəC sequences, it is possible that learners of other Slavic languages would be more lax in learning the proper coordination and acoustic cues for #CC since there are no #CəC phonotactic competitors. Whether or not this scenario is accurate is a matter for further research.

5 Conclusion

The phonotactics of Russian are uncommon in that sequences of obstruents have fewer positional restrictions than many other languages. As a consequence, Russian listeners and learners are not able to rely merely on phonotactic knowledge in order to perform speech segmentation and lexical acquisition. The results of this study have demonstrated that Russian speakers use durational cues to distinguish between consonant sequences that are phonotactically the same but differ in word boundary placement. This study also shows that the acoustic and articulatory patterns characteristic of Russian are not affected by English for the bilingual speakers. The durational cues apply to most types of consonant sequences, but can be overridden for other perceptual needs, such as ensuring that a stop release is audible. Finally, this study also shows that Russian speakers are consistent in their implementation of the coordination patterns for different types of sequences. Stop bursts for the first consonant are typically produced and schwas are almost never elided in #CC and C#C sequences. Since these acoustic cues reliably delineate between the types of consonant sequences investigated in this study, they should be available for various perceptual tasks carried out by Russian listeners and learners.

Acknowledgements

Thanks to Serafina Shishkova and Maria Gouskova for invaluable help with designing stimuli and recruiting and recording participants, and to two anonymous reviewers for their helpful comments. This work is supported by NSF CAREER grant BCS-0449560.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

Appendix A: Stimuli for spoken production

The following abbreviations are used in the tables below: #CC = cluster, #CəC = schwa, C#C = across word boundary. The notation (V) indicates the additional stimuli which were added to place a vowel in front of a fricative C1 (see section 2.3). The asterisk (*) indicates consonant combinations which were only read by monolinguals. The hyphen (-) indicates a morpheme boundary. Morpheme boundaries that are not near the sequences of interest are not indicated.

The UR column refers to the assumed underlying representation of each stimulus phrase, and the Surface form corresponds to the phonetic output (after vowel reduction, devoicing, voicing assimilation, etc.). The boxes in the Surface and UR columns indicate the target sequence.

Spaces indicated in the UR column indicates separation of orthographic words. In the case of sequence types other than C#C, this separation coincides with a preposition–word boundary within one phonological word. Devoicing indicated by the subscript (e.g. /v̥/) in the Surface column represents word-final devoicing, which is optional and/or gradient (see section 1.1 and references therein for discussion).

(A) /v/+Fricative

C1	C2	SEQUENCE	SURFACE	UR	GLOSS
v	z	#CC	'putʲ [vz] ɐr'val	'putʲ [v-z] o-r'val	he blew up the path
		#CC (V)*	kɐ'fe [vz] ɐr'val	ka'fe [v-z] o-r'val	(he) blew up the café
		#CəC	'zʲatʲ [vəz] ɐ'pil	'zʲatʲ [voz] -o'pil	the son-in-law cried out
		#CəC (V)*	u'ʒe [vəz] ɐ'pil	u'ʒe [voz] -o'pil	(he) already cried out
		C#C	'lʲe [v̥ z] ɐr'valsə	'lʲe [v z] ar'valsʲa	the lion showed off
v	ʒ	#CC	zɐ'lʲestʲ [vʒ] ɐ'kʲet	za'lʲestʲ [v ʒ] a'kʲet	to slip into the jacket
		#CC (V)*	ɐ 'tʲe [vʒ] ɐ'kʲet	a 'tʲe [v ʒ] a'kʲet	and those into the jacket
		#CəC	pɐj'matʲ [vəʒ] ɐ'ka	poʲ'matʲ [voʒ] a'ka	to catch the leader
		#CəC (V)*	vɐ'snʲe [vəʒ] ɐ'ka	vo 'snʲe [voʒ] a'ka	in the leader's dream
		C#C	pʲrʲi'pʲe [v̥ ʒ] ɐ'kʲejə	prʲi'pʲe [v ʒ] o'kʲeja	the chorus of the jockey

(B) /v/+Nasal

v n	#CC	'pʲitʲ [vn] ɐ'kladku	'pʲitʲ [v-n] a'kladku	to drink with sugar
	#CC (V)*	'kto [vn] ɐ'klad̥ku 'pʲjot	'kto [v-n] a'kladku 'pʲjot	who with sugar drinks
	#CəC	'vʲidʲitʲ [vən] ɐ'logʲiji	'vʲidʲetʲ [v an] a'logʲiji	to see in the analogy
	#CəC (V)*	ɐd'no [vən] ɐ'logʲiji	od'no [v an] a'logʲiji	one in an analogy
	C#C	'nʲeskəlʲkə 'slo [v̥] [n] ɐ'dʲeʒdɨ	'nʲeskolʲko 'slo [v n] a'dʲeʒdɨ	a few words of hope
v m	#CC	pʲe'tʃʲot [vm] ɐr'kofkʲi	pʲe'tʃʲot [v m] or'kovkʲe	s/he bakes in the carrot
	#CC (V)*	ɐd'nakə [vm] ɐr'kofkʲi	od'nako [v m] or'kovkʲe	however in carrot
	#CəC	pʲlʲisʲ'nʲot [vəm] ɐ'rʲettə	plʲes'nʲot [v am] a'rʲetto	he will splash into the Amaretto
	#CəC (V)*	'lʲibə [vəm] ɐ'rʲettə	'lʲibo [v am] a'rʲetto	either in Amaretto
	C#C	'sʲemʲ 'funtə [v̥] [m] ɐr'kofkʲi	'sʲemʲ 'funto [v m] or'kovkʲe	seven pounds of carrots

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

(C) /v/+Stop

v	d	#CC	krʲi'tʃatʲ	vd	ɐ'gon	krʲi'tʃatʲ	v-d o'gon	to yell in pursuit
		#CC (V)*	'mɨ jiˈʃʲo	vd	ɐ'gon	'mɨ jeˈʃʲo	v-d o'gon	we're still in pursuit
		#CəC	'zʲatʲ	vəd	ɐ'laz̥	'zʲatʲ	vod- o-'laz	the son-in-law is a scuba-diver
		#CəC (V)*	unʲi'vo	vəd	ɐ'laz̥	u nʲe'vo	vod- o-'laz	he has a scuba diver
		C#C	ivɐ'no	v̥ d	ɐ'vʲerʲil	iva'no	v d o'vʲerʲil	Ivanov trusted

(D) Fricative+/v/

z	v	#CC	'slɨʃɨt	zv	ɐ'nok	'slɨʃɨt	zv o'nok	s/he hears the bell
		#CC (V)*	'etə jiˈjo	zv	ɐ'nok	'eto jeˈjo	zv o'nok	that's her bell
		#CəC	u'kaʒɨt	zəv	ɐ'rot	u'kaʒɨt	za-v o'rot	s/he'll show the turn
		#CəC (V)*	'kto	zəv	ɐ'rot 'zdʲeləl	'kto	za-v o'rot 'sdʲelal	who made the turn
		C#C	prʲi'vʲo	z̥ v	ɐ'rotə	prʲi'vʲo	z v o'rota	(he) delivered the gates
s	v	#CC	i'mʲejit	sv	ɐ'bodu	i'mʲejet	sv o'bodu	s/he has freedom
		#CC (V)*	tʲi'bʲe	sv	ɐ'bodu 'daʲi	tʲe'bʲe	sv o'bodu 'daʲi	(they) gave you freedom
		#CəC	bʲi'rut	səv	ɐ'kupnə	bʲe'rut	so-v o-'kupno	they take jointly
		#CəC (V)*	'fsʲe	səv	ɐ'kupnə 'dʲeləjut	'vsʲe	so-v o-'kupno 'dʲelajut	everyone does jointly
		C#C	'vʲe	s v	ɐʃ'ʲi	v 'ʲe	s v oʃ'ʲi	they went into the woods

(E) Fricative+Fricative

f	s	#CC	dʲir'ʒatʲ	ʏs	ɐʏ'kʲe	dʲer'ʒatʲ	v s ov'kʲe	to hold in the trowel
		#CC (V)*	ɐ'nʲi	ʏs	ɐʏ'kʲe ɐ'staʲisʲ	o'nʲi	v s ov'kʲe o'staʲisʲ	they remained in the trowel
		#CəC	'tʲestʲ	fəs	ɐ'val	'tʲestʲ	fas o'val	the father-in-law pre-packed
		#CəC (V)*	vnu'trʲi	fəs	ɐ'val	vnu'trʲi	fas o'val	(he) pre-packed inside
		C#C	'mʲi	f s	ɐ'vɨ	'mʲi	f s o'vɨ	the myth of the owl

(F) Fricative+Nasal

z	n	#CC	'vrag	zn	ɐ'komɨj	'vrag	zn a'komɨj	familiar foe
		#CC (V)*	'novɨj 'no	zn	ɐ'komɨj	'novɨj 'no	zn a'komɨj	new, but familiar
		#CəC	dɐ'sʲtʲiɡ̥	zən	ɐ'vʲeskʲi	do'sʲtʲig	zan a'vʲeski	he reached/attained the curtain
		#CəC (V)*	'sʲnʲal jiˈvo	zən	ɐ'vʲesʲkʲi	'snʲal jeˈvo	za-n a-'vʲeskʲi	removed his curtains
		C#C	pʲrʲi'vʲo	z̥ n	ɐ'vʲernə: vʲi'no	prʲi'vʲo	z n a'vʲernoje vʲi'no	he probably brought wine
ʃ	m	#CC (V)*	kɐ'mu	ʃm	ɐ'tok	ko'mu	ʃm a'tok	to whom (should) the piece (go)
		#CəC (V)*	tɐ'bu	ʃəm	ɐ'nʲitʃʲiskəjə	ta'bu	ʃam a'nʲitʃʲeskoje	shamanistic taboo
		C#C*	'm	uʒ̥ m	ɐʃi'nʲist	'mu	ʒ m aʃi'nʲist	(my) husband is a machinist

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

(G) Fricative+Stop

z	b	#CC	'tak zb ɐl'tal	'tak s-b ol'tal	he shook it up this way
		#CC (V)*	'mnʲe zb ɐl'tal kɐk'tejlʲ	'mnʲe s-b ol'tal kok'tejlʲ	(he) mixed me a cocktail
		#CəC	'kak zəb ɐ'lʲel	'kak za-b o'lʲel	how he got sick
		#CəC (V)*	u'ʒe zəb ɐ'lʲel	u'ʒe za-b o'lʲel	(he) already got sick
		C#C	iʃʲ'ʃʲe z̬ b ɐlʲ'noj	iz-tʃʲe z b olʲ'noj	the sick man disappeared
z	d	#CC (V)*	'skorə zd ɐ'vatʲ	'skoro s-d a'vatʲ	soon to take
		#CəC (V)*	'mnogə zəd ɐ'vatʲ	'mnogo zad a'vatʲ	to give a lot
		C#C*	'obrə z̬ d ɐ'vʲerʲijə	'obra z d o'vʲerʲija	image of faith
z	g	#CC	fʲi'lʲip zg əvɐ'rʲil	fʲi'lʲip s-g ovo'rʲil	Phillip consented
		#CC (V)*	tɐg'da zg əvɐ'rʲil	tog'da s-g ovo'rʲil	(he) then consented
		#CəC	'trup zəg əvɐ'rʲil	'trup za-g ovo'rʲil	the dead body started speaking
		#CəC (V)*	i 'ja zəg əvɐ'rʲil	i 'ja za-g ovo'rʲil	and I started speaking
		C#C	zɐ'ka z̬ g ɐ'loʃ	za'ka z g o'loʃ	an order of galoshes
f	k	#CC	'klʲetʲ fk ɐ'tʲil	'klʲetʲ v-k a'tʲil	the cage rolled in
		#CC (V)*	ftʃʲi'ra fk ɐ'tʲil	vtʃʲe'ra v-k a'tʲil	rolled in yesterday
		#CəC	pəsʲi'titʲ fək ulʲ'tʲet	posʲe'titʲ fak ulʲtʲet	to visit the department
		#CəC (V)*	kɐg'da fək ulʲ'tʲet sɐz'dalsə	kog'da fak ulʲ'tʲet soz'dalsʲa	when the department was founded
		C#C	ʒɨ'ra f k ulʲ'turnɨj	ʒɨ'ra f k ulʲ'turnɨj	the giraffe is cultured
f	t	#CC	'sʲetʲ ft ɐ'ʃʲil	'sʲetʲ v-t a'ʃʲil	he dragged in the net
		#CC (V)*	vru'na ft ɐ'ʃʲil	vru'na v-t a'ʃʲil	(he) dragged the liar in
		#CəC	'ʒutʲ fət ɐ'lizm	'ʒutʲ fat a'lizm	a horror, fatalism
		#CəC (V)*	prʲi'tʃʲinə bɨ'la fət ɐ'lizm	prʲi'tʃʲina bɨ'la fat a'lizm	the reason was fatalism
		C#C	'gra f t ɐ'ʃʲil	'gra f t a'ʃʲil	the Count dragged (something)

(H) Stop+/v/

k	v	#CC	'xotʲ kv ɐx'tal	'xotʲ kv ox'tal	it at least clucked
		#CəC	u'znatʲ kəv ɐ'lʲerə	u'znatʲ kav a'lʲera	to recognize the cavalier
		C#C	'pois k v ɐ'rot	'pois k v a'rot	the search of the gates
d	v	#CC	kʲir'pʲitʃʲ dv ɐ'ra	kʲir'pʲitʃʲ dv o'ra	a brick of the courtyard
		#CəC	'pʲotr pʲi'trovʲitʃʲ dəv ɐ'lok	'pʲotr pʲe'trovʲitʃʲ do-v o'lok	Pyotr Petrovich dragged
		C#C	'rʲa d̬ v ɐ'rot	'rʲa d v o'rot	a row of gates
g	v	#CC	'pʲatʲ gv ɐzʲ'dʲik	'pʲatʲ gv ozʲ'dʲik	five carnations
		#CəC	'znatʲ gəv əru'na	'znatʲ gov oru'na	to know a talkative person
		C#C	'on pɐd'ʒo k v ɐ'rotə	'on pod'ʒo g v o'rota	he set fire to the gates

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

(I) Stop+Fricative

k	s	#CC*	xɐ'dʲitʲ k s ɐ'bakʲi	xo'dʲitʲ k s o'bakʲe	to walk towards the dog
		#CəC*	'tʃutʲ kəs ɐ'glazɨj	'tʃutʲ kos o'glazɨj	a little cross-eyed
		C#C*	utʃʲi'nʲi k s ɐ'glasʲin	utʃʲe'nʲi k s o'glasʲen	the student is in agreement
g	z	#CC*	dɐ'bavʲitʲ g z ɐr'platʲi	do'bavʲitʲ g z ar'platʲe	to add to (one's) pay
		#CəC*	unʲi'jo gəz ɐ'mʲer	u nʲe'jo gaz o'mʲer	she has a gas meter
		C#C*	pɐ'mo g z ɐ'mʲerʲitʲ	po'mo g z a'mʲerʲitʲ	(he) helped measure
p	ʃ	#CC	vɐ'rʲitʲ pʃ ɨ'no	va'rʲitʲ pʃ ɨ'no	to cook millet
		#CəC	'tʃutʲ pəʃ ɨvʲi'ɬitʲ	'tʃutʲ po-ʃ ɨvʲe'ɬitʲ	to stir a little
		C#C	u'to p ʃ ɨ'nok	u'to p ʃ ɨ'nok	the tavern was flooded
p	s	#CC	'pʲetʲ ps ɐ'lom	'pʲetʲ ps a'lom	to sing a psalm
		#CəC	'xotʲ pəs ɐ'dʲil	'xotʲ po-s a'dʲil	at least planted
		C#C	pɐ'to p s ɐ'lon	po'to p s a'lon	the salon flooded

(J) Stop+Nasal

g	n	#CC	'pusʲtʲ gn ɐ'jitsə	'pusʲtʲ gn o'jitsʲa	let it suppurate
		#CəC	zəplɐ'tʲitʲ gən ɐ'rar	zapla'tʲitʲ gon o'rar	to pay an honorarium
		C#C	trʲi'vo g̥ n ɐ'jitʲiji	trʲe'vo g n a'jitʲije	of the anxieties of the revelation
k	n	#CC	nʲi pɐdʲ'nʲatʲ kn u'ta	nʲe pod'nʲatʲ kn u'ta	not to raise a whip
		#CəC	pɐ'strojitʲ kən u'ru	po'strojitʲ kon u'ru	to build a kennel
		C#C	'ta k 'n udnə	'ta k 'n udno	it's so boring
k	m	#CC	dɐ'bavʲitʲ km ɐn'datu	do'bavʲitʲ k m anda'tu	to add to the mandate
		#CəC	ɬu'bʲitʲ kəm ɐnʲ'dʲirə	ɬu'bʲitʲ kom an'dʲira	to love the commander
		C#C	prʲivɐ'lo k m ɐn'dat	prʲivo'lo k m an'dat	he dragged over the mandate
p	n	#CC	ɐ'pʲatʲ 'pn u	o'pʲatʲ 'pn u	I'll spurn again
		#CəC	xɐ'tʲetʲ pən u'katʲ	xo'tʲetʲ po-n u'katʲ	to want to drive on
		C#C	'po p 'n udnɨj	'po p 'n udnɨj	the priest is boring
d	n	#CC	u'vʲidʲif 'dn o	u'vʲidʲiv 'dn o	having seen the bottom
		#CəC	'sʲemʲ stə'kanəy̯ dən ɐ'ʃu	'sʲemʲ sta'kanov do-n o'ʃu	I'll bring seven glasses
		C#C	'ɬo d̥ 'n otʃʲju 'tajit	'ɬo d 'n otʃʲju 'tajet	the ice melts at night

(K) Stop+Stop

t	k	#CC	'sʲemʲ 'fabrʲik tk ɐnʲ'ja	'sʲemʲ 'fabrʲik tk anʲ'ja	seven weaving factories
		#CəC	'kak tək ɐ'va	'kak tak o'va	as such
		C#C	nɐ'ɬi t k ɐ'nʲjak	na'ɬi t k o'nʲjak	the cognac is poured

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

k	t	#CC	pəvʲir'nutʲ [kt] ɐk'sʲi	povʲer'nutʲ [k t] ak'sʲi	to turn to the taxi
		#CəC	'sutʲ [kət] ɐ'strof	'sutʲ [kat] a'strof	the essence of the catastrophes
		C#C	du'ra [k t] ɐk'sʲist	du'ra [k t] ak'sʲist	the taxi driver is an idiot
k	p	#CC	'ʃasʲtʲ [kp] ɐ'rogu	'ʃastʲ [k p] o'rogu	walk quickly towards the threshold
		#CəC	pʲirʲi'statʲ [kəp] ɐ'ʃɨtʲ	pʲerʲe'statʲ [kop] a'ʃɨtʲ	to stop searching
		C#C	dɐ'lʲo [k p] ɐ'rom	da'lʲo [k p] a'rom	the ferry is far
g	d	#CC	'xotʲ [gd] ɐ'mam	'xotʲ [k d] o'mam	at least to the houses
		#CəC	prətʃʲi'tatʲ [gəd] ɐ'voj	protʃʲi'tatʲ [god] o'voj	to read the annual
		C#C	pɐ'bʲe [g̟ d] ɐ'moj	po'bʲe [g d] o'moj	the run (towards) home
g	b	#CC	'lʲesʲtʲ [gb] ɐ'ronu	'lʲestʲ [k b] a'ronu	to flatter the baron
		#CəC	ɐtsenʲ'itʲ [gəb] ɐ'rʲit	otse'nʲitʲ [gab] a'rʲit	to evaluate the size
		C#C	'dʲenʲi [g̟ b] ɐ'ronu	'dʲenʲe [g b] a'ronu	money for the baron

Appendix B: Sample spectrograms

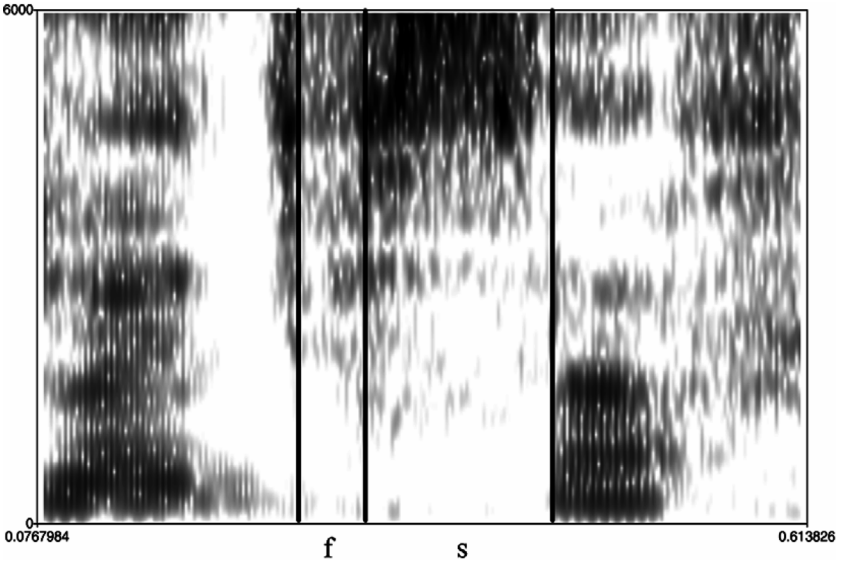


Figure B1 [#fs], following palatalized [tʲ], from the phrase [dʲirʒatʲ ɣsɐɣkʲe] 'to hold in the trowel' (Speaker 0814).

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

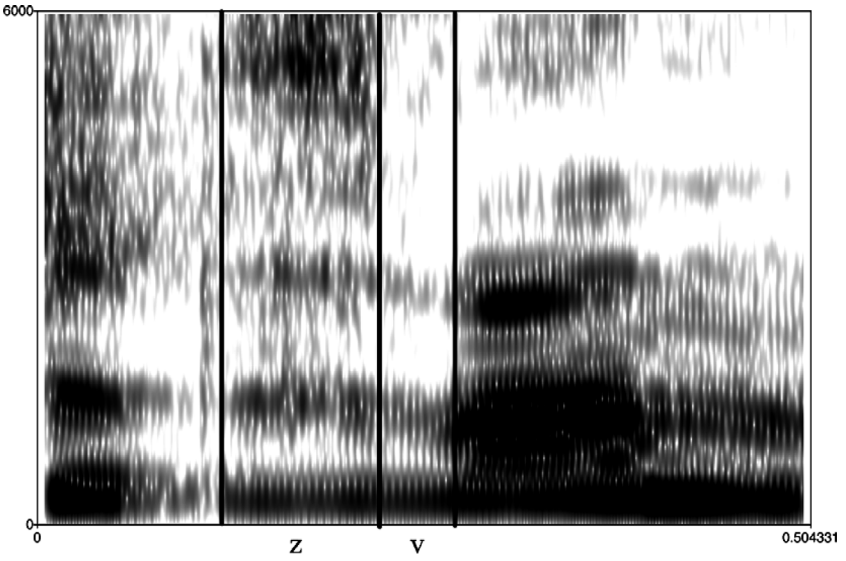


Figure B2 [#zv], following plain [t], from the phrase [slʲɪʃɨt **zvɐ**'nok] 's/he hears the bell' (Speaker 1110).

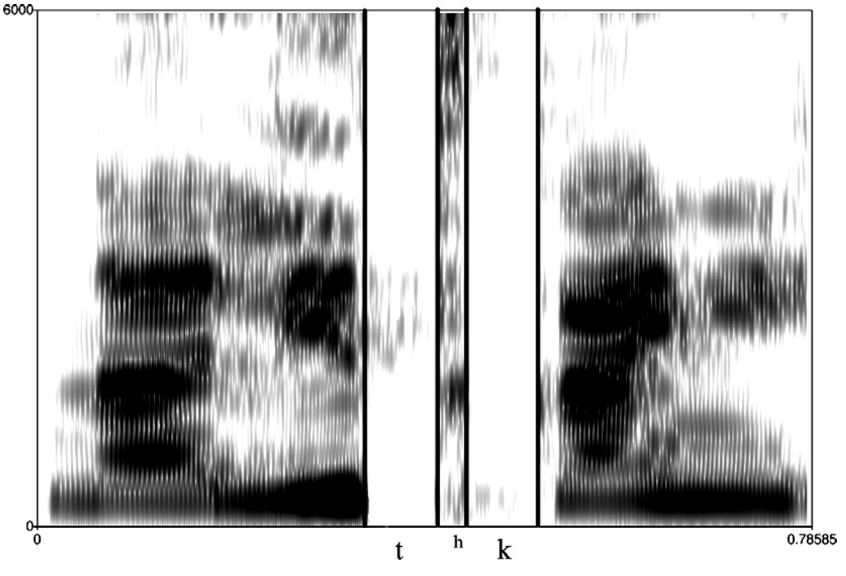


Figure B3 [t#k], from the phrase [nɐlʲʲɪt **kɐn**ʲjak] 'the cognac is poured' (Speaker 1110). The [ʰ] marks the burst of the [t].

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

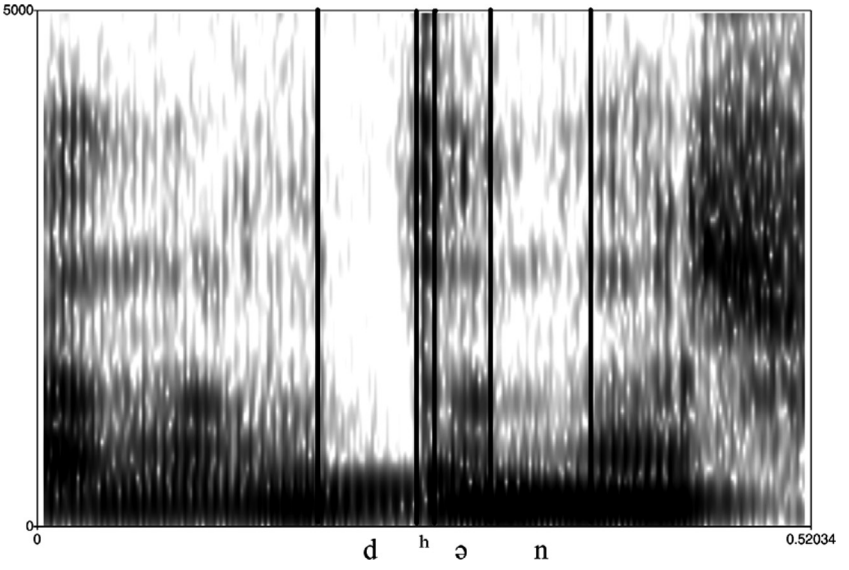


Figure B4 [dən], from the phrase [sʲemʲ stəkanəɣ **dənʁʃu**] 'I'll bring seven glasses' (Speaker 0610). The [ʰ] marks the burst of the [d].

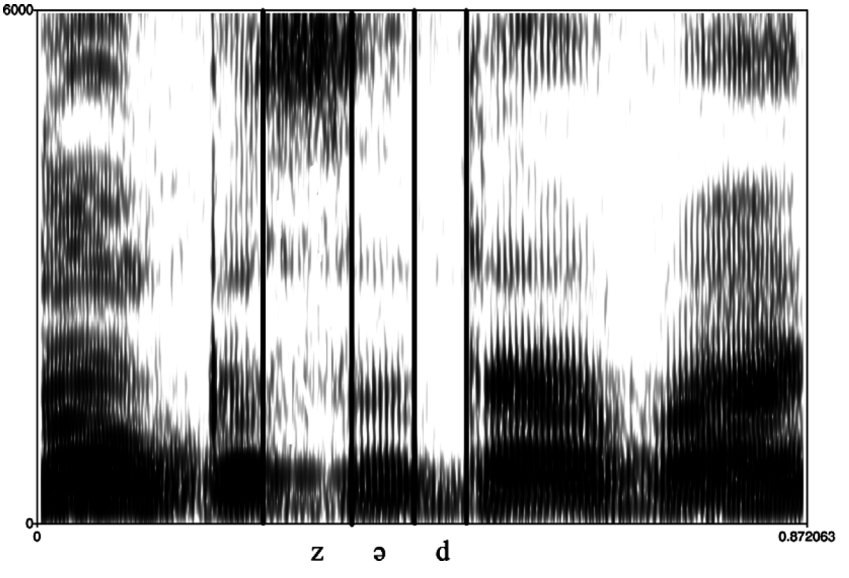


Figure B5 [zəd], from the phrase [mnogə **zədʁvatʲ**] 'to give a lot' (Speaker 1419).

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

References

Avanesov, Ruben Ivanovich. 1984. *Russkoje literaturnoje proiznoshenije* [Russian literary pronunciation]. Moscow: Prosveshchenije.

Boersma, Paul & David Weenink. 2006. Praat 4.3: Doing phonetics by computer. http://www.praat.org (13 February 2008).

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

Bradlow, Ann, David Pisoni, Reiko Akahane-Yamada & Yoh'ichi Tohkura. 1996. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America* 101, 2299–2310.

Browman, Catherine & Louis Goldstein. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica* 45, 140–155.

Browman, Catherine & Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In John Kingston & Mary Beckman (eds.), *Papers in Laboratory Phonology* I: *Between the grammar and physics of speech*, 341–376 Cambridge: Cambridge University Press.

Browman, Catherine & Louis Goldstein. 1995. Gestural syllable position effects in American English. In Fredericka Bell-Berti & Lawrence Raphael (eds.), *Producing speech: Contemporary issues for Katherine Safford Harris*, 19–33 New York: American Insititue of Physics.

Burton, Martha & Karen Robblee. 1997. A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics* 25, 97–114.

Byrd, Dani. 1996. Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24, 209–244.

Byrd, Dani. 2000. Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica* 57, 3–16.

Byrd, Dani & Cheng Cheng Tan. 1996. Saying consonant clusters quickly. *Journal of Phonetics* 24, 263–282.

Chitoran, Ioana. 1998. Georgian harmonic clusters: Phonetic cues to phonological representation. *Phonology* 15, 121–141.

Cho, Taehong & Patricia Keating. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29, 155–190.

Davidson, Lisa. 2005. Addressing phonological questions with ultrasound. *Clinical Linguistics and Phonetics* 19, 619–633.

Davidson, Lisa. 2006a. Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics* 34, 104–137.

Davidson, Lisa. 2006b. Schwa elision in fast speech: Segmental deletion or gestural overlap? *Phonetica* 63, 79–112.

Dmitrieva, Olga. 2005. Final devoicing in Russian: Acoustic evidence of incomplete neutralization. *Journal of the Acoustical Society of America* 117, 2570.

Elson, Benjamin. 1956. Sierra Popoluca syllable structure. *International Journal of American Linguistics* 13, 13–17.

Flege, James Emil & James Hillenbrand. 1984. Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America* 78, 708–721.

Fougeron, Cécile & Patricia Keating. 1997. Articulatory strengthening at the edges of prosodic domains. *Journal of the Acoustical Society of America* 101, 3728–3740.

Friederici, Angela & Jeanine Wessels. 1993. Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception & Psychophysics* 54, 287–295.

Gafos, Adamantios. 2002. A grammar of gestural coordination. *Natural Language & Linguistic Theory* 20, 269–337.

Gow, David & Peter C. Gordon. 1995. Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception & Performance* 21, 344–359.

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23, 387–429.

Halle, Morris. 1971. *The sound pattern of Russian*. The Hague & Paris: Mouton.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

Hamilton, William. 1980. *Introduction to Russian phonology and word structure*. Columbus, OH: Slavica Publishers.
Hayes, Bruce. 1984. The phonetics and phonology of Russian voicing assimilation. In Mark Aronoff & Richard Oehrle (eds.), *Language sound structure*, 318–328. Cambridge, MA: MIT Press.
Henderson, Janet & Bruno Repp. 1982. Is a stop consonant released when followed by another stop consonant? *Phonetica* 39, 71–82.
Hohne, Elizabeth & Peter Jusczyk. 1994. Two-month-old infants' sensitivity to allophonic differences. *Perception & Psychophysics* 56, 613–623.
Jones, Daniel & Dennis Ward. 1969. *The phonetics of Russian*. Cambridge: Cambridge University Press.
Jusczyk, Peter, Elizabeth Hohne & Angela Bauman. 1999. Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics* 61, 1465–1476.
Klatt, Dennis. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59, 1208–1220.
Lehiste, Ilse. 1960. An acoustic-phonetic study of internal open juncture. *Phonetica* Supplement 5.
Lehiste, Ilse. 1972. The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America* 35, 1773–1781.
Lindblom, Bjorn, Bertil Lyberg & Karin Holmgen. 1981. *Durational patterns of Swedish phonology: Do they reflect short-term memory processes?* Bloomington, IN: Indiana University Linguistics Club.
Luce, Paul & Edward A. Lyons. 1999. Processing lexically embedded words. *Journal of Experimental Psychology: Human Perception & Performance* 25, 174–183.
Lulich, Steven. 2004. Russian [v]: An acoustic study. *Folia Linguistica* 38, 63–85.
Matteson, Esther & Kenneth Pike. 1958. Non-phonemic transition vocoids in Piro (Arawak). *Miscellanea Phonetica* 3, 22–30.
Mattys, Sven, Laurence White & James Melhorn. 2005. Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General* 134, 477–500.
Mayo, Lynn Hansberry, Mary Florentine & Soren Buss. 1997. Age of second-language acquisition and perception of speech in noise. *Journal of Speech and Hearing Research* 40, 686–693.
McQueen, James. 1998. Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39, 21–46.
Morelli, Frida. 1999. *The phonotactics and phonology of obstruent clusters in Optimality Theory*. Ph.D. dissertation, University of Maryland.
Norris, Dennis. 1994. Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52, 189–234.
Norris, Dennis, James McQueen, Anne Cutler & Sally Butterfield. 1997. The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology* 34, 191–243.
Oller, D. Kimbrough. 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54, 1235–1247.
Ozhegov, Sergei Ivanovich. 1990. *Slovar' russkogo jazyka* [Dictionary of the Russian language]. Moscow: Russkij jazyk.
Padgett, Jaye. To appear. Russian voicing assimilation, final devoicing, and the problem of [v]. *Natural Language & Linguistic Theory*.
Padgett, Jaye & Marija Tabain. 2005. Adaptive Dispersion Theory and phonological vowel reduction in Russian. *Phonetica* 62, 14–54.
Pye, Shizuka. 1986. Word-final devoicing of obstruents in Russian. In Sarah Hawkins (ed.), *Cambridge papers in phonetics and experimental linguistics*, vol. 5, 1–10. Cambridge: Cambridge University Press.
Quené, Hugo. 1992. Durational cues for word segmentation in Dutch. *Journal of Phonetics* 20, 331–350.
Salverda, Anne Pier, Delphine Dahan & James McQueen. 2003. The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition* 90, 51–89.
Sancier, Michele & Carol Fowler. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25, 421–436.
Scott, Donia. 1982. Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America* 71, 996–1007.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press

Solé, Maria Josep. 1997. Timing patterns in the production of a foreign language. In Lourdes Díaz & Carmela Pérez (eds.), *Views on the acquisition and use of a second language: The 7th Annual European Second Language Association Conference*, 539–551. Barcelona: Universitat Pompeu Fabra.

Tajima, Keiichi, Robert Port & Jonathan Dalby. 1997. Effects of termoral correction on intelligibility of foreign accented English. *Journal of Phonetics* 25, 1–24.

Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.

Weber, Andrea. 2000. Phonotactic and acoustic cues for word segmentation in English. In B. Yuan, T. Huang & X. Tang (eds.), *The 6th International Conference on Spoken Language Processing*, vol. 3, 782–785. Beijing.

Wright, Richard. 1996. *Consonant clusters and cue preservation in Tsou*. Ph.D. dissertation, UCLA.

Yanagawa, Mariko. 2007. Consonant timing in L2 English: The emergence of a default pattern. In Jennifer Cole & José Ignacio Hualde (eds.), *Laboratory Phonology* 9. Berlin: Mouton de Gruyter.

Zsiga, Elizabeth. 1994. Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics* 22, 121–140.

Zsiga, Elizabeth. 2003. Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition* 25, 399–432.

https://doi.org/10.1017/S0025100308003447 Published online by Cambridge University Press