

# Acoustics of Russian voiceless sibilant fricatives

Alexei Kochetov

Department of Linguistics, University of Toronto

[al.kochetov@utoronto.ca](mailto:al.kochetov@utoronto.ca)

This study investigated acoustic properties of the four-way contrast in Russian voiceless sibilant fricatives (/s s<sup>j</sup> ʂ ʂ<sup>j</sup>/). Words with these consonants, occurring in a variety of phonetic contexts, were elicited from 10 Russian native speakers. Measurements were made of duration, centre of gravity (COG) and intensity of fricative noise, as well as of formants F1–F3 during the following vowel. The results revealed that the anterior vs. posterior contrast was clearly distinguished by COG throughout the frication period, and to a lesser degree by noise intensity. The palatalized vs. non-palatalized contrast was distinguished by F1 and (especially) F2 at the onset, the midpoint, and, in some cases, at the offset of the following vowel. The four-way contrast was thus well categorized by a combination of COG and F2–F1 differences. Fricative duration only marginally distinguished /ʂ<sup>j</sup>/, commonly described as geminate, from the other consonants. Clear durational differences, however, were observed for the same consonants in different positional and stress contexts. Overall, the results of the study provide a systematic documentation of the typologically uncommon fricative contrast, while also shedding light on some facts of synchronic patterning and historical development of the fricative system of Russian.

## 1 Introduction

Cross-linguistic surveys of fricative systems (Maddieson 1984, Ladefoged & Maddieson 1996) show that languages rarely contrast more than two phonemic sibilant fricatives of the same phonation type. Even less common is the use of secondary articulations to contrast sibilants within the same primary place. In Maddieson's (1984) survey of sound inventories, only five out of 317 languages have a palatalized dental/alveolar /s<sup>j</sup>/ (compared to 276 languages with a plain dental/alveolar /s/), and only three languages have a palatalized post-alveolar /ʂ<sup>j</sup>/ (compared to 163 languages with a non-palatalized posterior sibilant). These small numbers can be explained by the relative synchronic/diachronic instability of secondary articulation contrasts. The articulatorily complex /s<sup>j</sup>/ (involving the partly conflicting tongue tip and tongue body gestures) is prone to shifting in place to /ʃ/ or /ç/, as part of a more general place-changing palatalization of coronals before front vowels and /j/ (Bhat 1978, among others). The posterior /ʂ<sup>j</sup>/, on the other hand, does not differ considerably from /ʃ/ or /ç/ (which are inherently more or less palatalized), and thus is often subject to merger or depalatalization (see Carlton 1990 on palatalization and depalatalization processes in Slavic). Interestingly, the few languages that do have phonemic /s<sup>j</sup>/ and/or /ʂ<sup>j</sup>/ (Bulgarian, Kashmiri, Ket, Lithuanian, Paez, Russian, and Yurak in Maddieson 1984) have those segments as part of a robust palatalized vs. non-palatalized contrast involving consonants of other manners and places. This suggests that the historical stability of these consonants is at least

**Table 1** Sample words with Russian voiceless sibilant fricatives (in bold) in various positions.

Position	Anterior ('hissing')		Posterior ('hushing')	
	Non-palatalized	Palatalized	Non-palatalized (retroflex)	Palatalized (prepalatal)
Word-initial	<b>[sok]</b> 'juice'	<b>[sʲok]</b> 'whipped'	<b>[ʂok]</b> 'shock'	<b>[ʃʲok]</b> 'cheeks, GEN'
Word-medial	<b>[unʲi'son]</b> 'unison'	<b>[unʲi'sʲon]</b> 'carried away'	<b>[razrʲi'son]</b> 'allowed'	<b>[razmʲi'ʃʲon]</b> 'placed'
Word-final	<b>[zbroʂ]</b> 'reset'	<b>[broʂʲ]</b> 'drop it'	<b>[broʂ]</b> 'brooch'	<b>[xvoʂʲ]</b> 'equisetum'

in part contingent on their integration into a broader system of phonemic contrasts. Another potentially important factor is the sufficient acoustic differentiation of palatalized and non-palatalized anterior and posterior sibilants – the question that is specifically investigated in this paper.

Russian is one of the few languages that have contrastive palatalized sibilant fricatives. These consonants – the dental/alveolar /sʲ/ and the palatalized post-alveolar (prepalatal) /ʃʲ/ – are contrastive with the non-palatalized dental/alveolar /s/ and the retroflex (apical post-alveolar) /ʂ/ (Timberlake 2004).<sup>1</sup> The secondary articulation contrast intersects with the primary place contrast between the anterior (also called 'hissing') and posterior ('hushing') coronal fricatives. Sample words with the four-way contrast among voiceless fricatives in initial, medial, and final position are shown in Table 1.

Anterior fricatives are well-integrated into the Russian consonant inventory, being part of the phonemic non-palatalized vs. palatalized distinction (which is particularly robust among anterior coronals, i.e. /t d s z n l r/ vs. /tʲ dʲ sʲ zʲ nʲ lʲ rʲ/; Timberlake 2004). In contrast, the status of posterior fricatives within the palatalization correlation (set of contrasts) is less clear. Partly this is due to phonetic characteristics of these consonants other than secondary articulation (Jones & Ward 1969, Matusевич 1976, Avanesov 1984). For example, /ʂ/ is characterized by a retroflex-like (tip-up) primary constriction, a retraction of the tongue back (see Keating 1991, Hamann 2004), and some degree of lip rounding (Jones & Ward 1969, Matusевич 1976; but see Ladefoged & Maddieson 1996: 148, who attributed no lip rounding to the sound). In contrast, the tongue tip for /ʃʲ/ is lowered down, with the constriction made with the blade in the prepalatal region; the lips are also rounded, and the consonant has overall longer duration (i.e. [ʃʲː]) (Bolla 1981). This length can be attributed to the historical development of this consonant from a sequence /s/ + /tʃʲ/ (via [ʃʲ tʃʲ]); Matusевич 1976, Timberlake 2004).<sup>2</sup> Given the marginal role of consonant length in Russian,<sup>3</sup> some researchers in the past have disputed the phonemic status of /ʃʲ/, still considering it an underlying sequence of distinct phonemes (/s/ and /tʃʲ/; Baudouin de Courtenay 1964, Zinder 1963; but see Scherba 1957, Avanesov 1948, cited in Matusевич 1976; see also Timberlake 2004). Another reason for the special treatment of the posterior fricatives /ʂ/ and /ʃʲ/ is their distinct phonological behavior. These segments fail to alternate with each other in contexts where other non-palatalized and palatalized consonants do (e.g. [duʂa] – [duʂe], \*[duʃʲe] 'soul, NOM.SG – GEN.SG; compare [kasa] – [kasʲe] 'plait, NOM.SG – GEN.SG). Given this, Timberlake (2004: 57) refers to the posterior fricatives as 'immutable hard and soft'

<sup>1</sup> All four voiceless fricatives have voiced counterparts /z zʲ zʂ zʃʲ/, although the latter is phonemically marginal, occurring only in a handful of words (Timberlake 2004: 66).

<sup>2</sup> In addition to /ʃʲ/ that is phonemic (e.g. in [borʃʲ] 'borsht'), Russian has [ʃʲ] derived from word-internal sequences of /s z ʂ zʲ + /tʃʲ/ across morpheme boundaries (e.g. [ra'skaʃʲik] from /raz-skaz-tʃʲik/ (Timberlake 2004: 65–66).

<sup>3</sup> Most Russian consonants have geminate counterparts; however, these are largely limited to loanwords or certain morphological contexts (see Timberlake 2004: 67–68; Dmitrieva 2017).

(non-palatalized and palatalized) sounds, in contrast to the ‘mutable’ /s/ and /sʲ/, among other consonants.

While there have been a number of articulatory studies of Russian sibilant fricatives (Skalozub 1963, Bolla 1981, Kedrova et al. 2008, Litvin 2014, among others), relatively little has been done on their acoustic properties. A few studies that did explore the acoustics of Russian fricatives were often limited to a subset of the contrasts or vowel contexts, and/or did not always examine acoustic differences quantitatively (Fant 1960, Shupljakov, Fant & de Serpa-Leitao 1968, Derkach, Fant & de Serpa-Leitao 1970, Bolla 1981, Zsiga 2000, Padgett & Żygis 2007). The goal of this study is to provide a systematic investigation of durational and spectral properties distinguishing the Russian four-way contrast in voiceless sibilant fricatives. This is done by examining a number of acoustic variables (consonant duration, frication noise intensity and frequency, as well as adjacent vowel duration and formants F1–F3) in the four fricatives occurring before five vowels and in four positional/stress conditions. Taken more broadly, the investigation of Russian fricatives is intended to contribute to phonetic documentation of cross-linguistically rare fricative contrasts (Ladefoged & Maddieson 1996, Gordon, Barthmaier & Sands 2002, among others).

## 2 Acoustic properties of Russian fricatives: Previous work

### 2.1 Duration

Although many descriptive phonetic studies mention length as a distinctive property of /ʃʲ/, surprisingly little quantitative data exist on the duration properties of fricatives in Russian. Bolla’s (1981) acoustic description of various Russian consonants (based on a single male speaker) includes the following duration measurements for voiceless sibilant fricatives: 195 ms for /s/, 177 ms for /sʲ/, 187 ms for /ʃ/, and 290 ms for /ʃʲ/. Thus, the palatalized posterior fricative in his data is 1.55 times longer than its non-palatalized counterpart. It should be noted that the phonetic contexts in which the fricatives occur are not identical, and therefore, it is not clear to what extent these measurements are representative of the contrast in general. Duration measurements were also performed by Kochetov & Radišić (2009), who examined the production of four fricatives in [aʰCa]-type nonsense words by four speakers. The mean duration values for the fricatives were 180 ms for /s/, 196 ms for /sʲ/, 180 ms for /ʃ/, and 222 ms for /ʃʲ/, thus giving a ratio of 1.23 for the latter two consonants. Notably, not all the speakers (three out of four) significantly distinguished the contrasts between /ʃʲ/ and some of the other fricatives (/ʃ/ and /s/, but not /sʲ/). This suggested that the durational distinction is not as clear as previously reported, or perhaps has diminished over time. It is also possible, however, that these results are dialect-specific, given that all four speakers in the study were from the same region in Russia (the city of Perm).

Regardless of segment-specific length differences, consonants as a class can differ in duration depending on contextual factors such as syllable/word position and stress. For example, Kochetov & Lobanova (2007) found that Komi-Permyak sibilant fricatives were significantly shorter word-finally than word-initially, while Zsiga (2000) reported significant shortening of Russian stops in unstressed syllables. These contextual factors may well interact with segment-specific length differences (e.g. /ʃʲ/ vs. /ʃ/ in onset vs. coda); however, such interactions have not been investigated for Russian fricatives.

### 2.2 Fricative spectra: Intensity and frequency

Bolla (1981) observed that palatalized fricatives /sʲ/ and /ʃʲ/ showed higher intensity noise compared to their non-palatalized counterparts. No comparison was made between anterior

and posterior fricatives. Kochetov & Radišić (2009), however, did not find intensity to distinguish palatalized and non-palatalized fricatives. Instead, intensity measurements in their data differentiated anterior and posterior fricatives, with /s/ and /sʲ/ being significantly less intense than /ʃ/ and /ʃʲ/.

Bolla's (1981) informal examination of fricative spectra revealed higher concentrations of energy for anterior fricatives (above 4000 Hz with peaks around 5000–6000 Hz for /s/ and above 4500 Hz with peaks between 4750 Hz and 6000 Hz for /sʲ/) compared to posterior fricatives (above 3500 Hz with the main peak at 4500–5000 Hz for /ʃ/ and from overall, above 2500 Hz, main peak at around 3000 Hz). These noise distribution patterns are expected, given the 'hissing' and 'hushing' auditory classification of the sounds corresponding to the more anterior and more posterior sources of noise, respectively (see Hughes & Halle 1956, Jones & Ward 1969). However, specific details of fricative spectra should be taken with caution, given that the contexts for the fricatives were not matched.

Zsiga's (2000) study examined Russian /s ʃ sʲ/ in the VC(#)V context in real words produced by five speakers (four males and one female). Measurements of the centre of gravity (COG; also called 'centroid') of fricative noise were performed at the onset, midpoint, and offset of each fricative. Zsiga reported a midpoint COG of about 6000 Hz for /s/ and /sʲ/, and COG of below 5000 Hz for /ʃ/. These values were overall similar to COG values for English fricatives /s/ (5900 Hz) and /ʃ/ (5100 Hz) produced by five English speakers. Interestingly, however, Russian /sʲ/ exhibited little or no change in COG over time, compared to the English /s/ + /j/ sequence, where COG was gradually lowering.

Padgett & Žygis (2007) measured COG of four Russian voiceless fricatives (also at three points) produced by four speakers in nonsense CV syllables, as part of a larger study of Slavic fricative contrasts. They found that /s/ and /sʲ/ had COG around 6000 Hz (but somewhat lower for /sʲ/, especially at the fricative offset), while /ʃ/ and /ʃʲ/ had COG around 3000 Hz. These findings are similar to Kochetov & Radišić (2009), who reported an average midpoint COG of 5870 Hz for /s/, 5520 Hz for /sʲ/, 3705 Hz for /ʃ/, and 3845 Hz for /ʃʲ/ produced before /a/ by their four speakers (two males and two females). Significant differences in their results involved only pairs of anterior and posterior fricatives (consistently for all four speakers). One other study that examined Russian fricative spectra is Funatsu & Kiritani (1998), whose purpose was to investigate acoustic and perceptual differences between Russian and Japanese fricatives. The authors' data, based on Russian /s sʲ ʃ/ produced by three native speakers word-initially before /a o u/, showed a binary distribution consistent with other studies: COG at the fricative midpoint was higher for the anterior /s/ and /sʲ/ (4400 Hz and 4300 Hz) and lower for the posterior /ʃ/ (2300 Hz).

### 2.3 CV formant transitions

It has been long observed that palatalized consonants as a class, regardless of their manner and primary place of articulation, are cued by formant transitions from and to the adjacent vowels (Fant 1960, Shupljakov et al. 1968, Derkach et al. 1970, Purcell 1979). Specifically, these studies determined that in the same phonetic context, palatalized consonants are associated with considerably higher F2 and somewhat lower F1 than non-palatalized ones, especially at the onset of the following vowel. These differences are the result of fronting and raising of the tongue body for secondary palatalization, timed at the release of the consonant constriction. A measure that combines the effects of the first two formants, F2–F1 difference, has been useful in capturing differences between palatalized vs. non-palatalized contrasts (Iskarous & Kavitskaya 2010). Purcell's (1979) examination of dental/alveolar palatalized and non-palatalized stops /d/ and /dʲ/ (among other stops) produced by four speakers (two males and two females) in nonsense VCV words revealed average F1 and F2 differences of 82 Hz and 409 Hz, respectively, giving an F2–F1 difference of 1884 Hz for palatalized and 1393 Hz for non-palatalized coronal consonants. Similar findings were

obtained by Shupljakov et al. (1968) and Derkach et al. (1970), whose stimuli included sibilant fricatives.

Focusing specifically on fricatives, Funatsu & Kiritani (1998) measured F2 at the onset of postconsonantal vowels /a/, /o/, and /u/. They found that F2 was considerably higher after the palatalized /sʲ/ than the non-palatalized /s ʂ/, with the magnitude of this difference affected by the vowel context. Combined with COG of the fricative, F2 could distinguish three fricatives in three vowel contexts (with /ʃ/ not considered by the authors). The study by Kochetov & Radišić (2009) examined the effect of the four fricatives on F1 through F3 of adjacent vowels. They found that /sʲ/ and /ʃ/ had a lower F1 (490 Hz and 545 Hz) and higher F2 (1950 Hz and 1975 Hz) at the onset of the following vowel /a/, compared to /s/ and /ʂ/ (F1: 634 Hz and 630 Hz; F2: 1550 Hz and 1515 Hz). This gave F2–F1 differences of around 1450 Hz for palatalized and around 900 Hz for non-palatalized fricatives. Differences in F3 were not consistent. Overall, this suggests that both anterior and posterior fricatives in Russian have transitions similar to other palatalized and non-palatalized consonants, and thus appear to be strongly integrated into the phonemic system. Again, however, the latter findings with respect to fricatives were based on a single vowel context, and thus remain to be confirmed.

## 2.4 Summary

Table 2 presents a summary of previous acoustic studies of Russian sibilant fricatives, including numbers of speakers, examined contrasts and contexts, acoustic measurements, and major findings.<sup>4</sup> It can be seen that, although Russian fricatives have received a fair share of attention in the phonetic literature, few studies have examined fricative spectra and vowel transitions together, and even fewer have combined either of these measurements with those of consonant duration. Moreover, none of the studies have examined fricative realizations across positional or stress contexts. Taken together, however, the results suggest that Russian sibilant fricatives are distinguished by a combination of fricative spectra (the anterior vs. posterior contrast) and formant vowel transitions (the secondary articulation contrast). Spectral differences within the anterior and posterior categories appear to be smaller, and often not significant. Further, there is evidence that /ʃ/ may not be much longer than the other fricatives, contrary to the common assumption in the descriptive phonetic literature on Russian.

## 3 Russian sibilant fricatives: Phonotactics and historical development

The phonetic documentation of Russian fricatives is also expected to provide some insights into the phonological patterning and historical development of these sounds. While fricatives in Modern Standard Russian can occur in a range of contexts (as noted in Section 1 above; see Table 1), the palatalized fricatives /sʲ/ and /ʃ/ are in general more restricted than their non-palatalized counterparts /s/ and /ʂ/. Specifically, Table 3 (based on the discussion in Timberlake 2004: 58–60)<sup>5</sup> illustrates these distributional differences across four consonant cluster contexts – word-medially post- and preconsonantly (C\_V and V\_C), word-initially preconsonantly (#\_C), and word-finally postconsonantly (C\_#). All four fricatives can occur in the first context (C\_V). In the second context (V\_C), /sʲ/ and /ʃ/ can occur only before

<sup>4</sup> Both vowels in VCV sequences in Shupljakov et al. (1968) and Derkach et al. (1970) appeared to be pronounced as stressed. Gender of the speakers was not indicated in Padgett & Žygis (2007).

<sup>5</sup> The distribution of fricatives can be further affected by the secondary articulation of the adjacent consonant. If the following consonant is a palatalized coronal, /s/ is subject to regressive palatalization (changing to [sʲ]), depending on certain morphological and lexical factors (Avanesov 1984, Timberlake 2004).

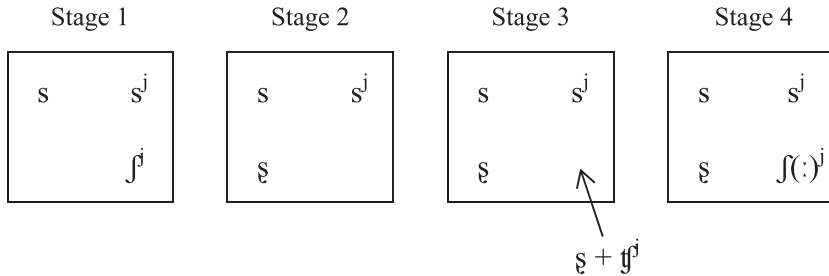
**Table 2** A summary of previous acoustic studies of Russian fricatives. '>' stands for 'greater than'; f = female, m = male.

Study	Speakers	Fricatives, context	Measurements	Relevant findings
Shupljakov et al. 1968	5 (3 f, 2 m)	/s sʲ ʂ ʃ/ (among other consonants) in symmetric [V_V] with five vowels	F2 and F3 at V1/V2 onset and offset	F2: sʲ, ʃʲ > s, ʂ F3: sʲ, ʃʲ > s, ʂ [V_V <sub>front</sub> ]
Derkach et al. 1970	1 (m)	/s sʲ ʂ ʃ/ (among other fricatives) in (a-)symmetric [V_V] with 5 vowels	F1, F2 and F3 at V1/V2 onset and offset	F1: s, ʂ > sʲ, ʃʲ F2: sʲ, ʃʲ > s, ʂ F3: sʲ, ʃʲ > s, ʂ
Bolla 1981	1 (m)	/s sʲ ʂ ʃ/ (among other consonants), not controlled for vowel context or position	Consonant duration Noise intensity (informal) Noise frequency (informal)	ʃʲ > ʂ sʲ, ʃʲ > s, ʂ s, sʲ > ʂ, ʃʲ
Funatsu & Kiritani 1998	3 (m)	/s sʲ ʂ/ in [#¹_V] with /a o u/	Noise frequency (COG at midpoint) F2 at V onset	s, sʲ > ʂ, ʃʲ sʲ > s, ʂ
Zsiga 2000	5 (4 f, 1 m)	/s sʲ ʂ/ in [¹a_(#)a] or [o_(#)a]	Noise frequency (COG at midpoint)	s, sʲ > ʂ
Padgett & Žygis 2007	4	/s sʲ ʂ ʃ/ in [¹a]	Noise frequency (COG at onset, midpoint, and offset)	s, sʲ > ʂ, ʃʲ
Kochetov & Radišić 2009	4 (2 f, 2 m)	/s sʲ ʂ ʃ/ in [¹a¹_a]	Consonant duration Noise intensity (at midpoint) Noise frequency (COG at midpoint) F1, F2, and F3 at following V onset	ʃʲ > ʂ ʂ, ʃʲ > s, sʲ s, sʲ > ʂ, ʃʲ F1: s, ʂ > sʲ, ʃʲ F2: sʲ, ʃʲ > s, ʂ

**Table 3** Phonotactic distribution of Russian voiceless fricatives in clusters. Yes = present, no = absent. C<sub>het</sub> = hetero-organic consonant, C<sub>hom</sub> = homorganic consonant, C<sub>obs</sub> = obstruent; dark and light shading indicates, respectively, more and less restricted patterns.

Context	/s/	/sʲ/	/ʂ/	/ʃʲ/	Examples
C_V	yes	yes	yes	yes	af'sa 'oats, GEN.SG', af'sʲanka 'oats porridge', lap'sa 'noodles', saap'ʃʲa 'together'
V_C	yes	yes/no	yes	yes/no	'pasmurnij 'cloudy', 'krasnij 'red', vasʲ'moj 'eighth', 'jašma 'jasper', 'strašnij 'scary', avajʲ'noj 'vegetable, ADJ' (but *sʲC <sub>het</sub> -, *ʃʲC <sub>het</sub> -)
C_#	yes	no	yes	yes/no	vors 'lint', gips 'plaster', farš 'ground meat', kofš 'ladle', boršʲ 'borst' (but *-Csʲ, *-C <sub>obs</sub> ʲ)
#_C	yes	no	yes	no	'spal 'he slept', 'spala 'sleepers' (but *sʲC-, *ʃʲC-)





**Figure 1** A diagram outlining stages in the historical development of Russian voiceless sibilant fricatives.

either hetero-organic or homorganic consonants, respectively. In the third context (C\_#), /sʲ/ is completely absent, while /ʃʲ/ can occur only after a sonorant consonant. Finally, in the fourth context (#\_C), neither of the palatalized fricatives can occur. Notably, none of these restrictions apply to the non-palatalized fricatives. The observed restrictions on palatalized sibilants may well be due to the way these consonants are acoustically cued. If palatalized sibilants are primarily distinguished by contextual cues such as formant transitions into the following vowel, these consonants would be disfavoured in contexts that lack this vowel – namely the V\_C, C\_#, and #\_C contexts (see Kochetov 2006 on Russian palatalized stops). It is therefore of interest to understand both consonant-internal (fricative noise frequency and intensity) and contextual (vowel transitions) acoustic properties of Russian fricatives.

Another question of interest is the historical development of Russian and Slavic sibilant systems (Hamann 2004, Padgett & Žygis 2007, Žygis & Padgett 2010). As Padgett & Žygis (2007) note, Russian sibilant fricatives have undergone a series of phonetic changes which can be best viewed as perceptual optimization through dispersion (Liljencrants & Lindblom 1972, Lindblom & Maddieson 1988, Flemming 2002). This process is schematically represented in Figure 1, based on the discussion in Borkovskii & Kuznetsov (1965). At Stage 1 (Old Russian, prior to the 12th century), the language exhibited a three-way sibilant contrast among the non-palatalized dental /s/, palatalized dental /sʲ/, and a palatalized post-alveolar or alveopalatal /ʃʲ/. At Stage 2 (by the 14th century), the latter consonant became retroflex (or apical post-alveolar), thus shifting away from /sʲ/. The resulting gap in the sibilant inventory was filled with a sequence /ɕ + ʧʲ/ at Stage 3. This sequence was originally produced as [ʃʲʧʲ], subsequently developing at Stage 4 into a new phoneme /ɕ(ː)ʲ/, which was inherently long ([ɕːʲ]). Presumably the shift of the original /ʃʲ/ to /ɕ/ was driven by the need to perceptually differentiate the former from the palatalized /sʲ/, which is potentially unstable (Hamann 2004, Padgett & Žygis 2007). The addition of the fourth fricative to the system, /ɕ(ː)ʲ/, could have served to stabilize the already well-dispersed sibilant contrasts, while extending the set of non-palatalized vs. palatalized contrasts to posterior coronals (i.e. re-using the already highly functional articulatory mechanism; see Clements 2003 on ‘feature economy’). In other words, the historical development of Russian sibilant fricatives possibly presents an interesting test case for Lindblom & Maddieson’s (1988: 72) proposed phonetic universal: ‘Consonant inventories tend to evolve so as to achieve maximal perceptual distinctiveness at minimum articulatory cost’.<sup>6</sup> Whether this interpretation of the historical development of Russian fricatives and its relevance to universal phonetic principles are correct depends to a large degree on our understanding of acoustic properties differentiating the four-way fricative contrast.

In sum, this paper is motivated by the need to acoustically document a typologically uncommon sibilant contrast, while also seeking answers to the long-standing questions of phonological patterning and sound change involving Russian fricatives.

<sup>6</sup> I am thankful to Shigeto Kawahara for bringing this issue and the paper to my attention.

## 4 Method

### 4.1 Speakers

The participants were 10 native speakers of Standard Russian, five females and five males. They were all born and raised in Russia<sup>7</sup> at least until the age of 16, and had lived in Canada for no more than six years. Their age ranged from 19 to 29 years, with a median of 21.5 years. At the time of the experiment, the participants were University of Toronto undergraduate or graduate students (none of whom had taken linguistics courses). While fluent in English, the participants reported to be using Russian on a day-to-day basis. They also reported to have normal speech and hearing.

### 4.2 Materials and the procedure

There were three sets of materials, all of which included words with the four fricatives /s sʲ ʃ ʃʲ/. In the first (and the largest) set, these consonants appeared word-initially and word-medially before the stressed vowels /a e i o u/.<sup>8</sup> For example, the set of words with these consonants in word-initial position before /o/ included [sok] ‘juice’, [sʲok] ‘whipped’, [ʃok] ‘shock’, and [ʃʲok] ‘cheeks, GEN’. There were a total of 40 words in this set, a complete list of which is given in Table A1 in the appendix. The purpose of this set was to systematically investigate the four-way contrast and its realization in two prevocalic positions in a variety of vowel contexts. Set 2 was designed to examine differences between the same consonants word-initially and word-finally. It included eight words (four of which were from Set 1) with four fricatives at the beginning or at the end of words, all either before or after a stressed /o/. The word-final context is also interesting because of the previously noted optional degemination of /ʃʲ/ (Avanesov 1984, Timberlake 2004). Set 3 was designed to explore the effect of stress, and included eight words where the same consonants appeared before stressed and unstressed /a/. Four of these words were from Set 1. Both word lists are given in Tables A2 and A3 in the appendix. All the items were randomized, embedded in a carrier phrase [ˈpapa \_\_\_ paftaˈrʲil] ‘Dad repeated \_\_\_.’ and presented in the Cyrillic script. Three repetitions of each item were recorded, giving 144 items per speaker (48 items × 3 repetitions).

Audio recordings were performed in a soundproof booth using a digital audio recorder Fostex FR-2 and an ATR3035 condenser microphone with the quantization of 16 bits and the sampling rate of 44000 Hz. Altogether, 1321 tokens (the intended 1320 items minus nine omissions or reading errors) were collected.

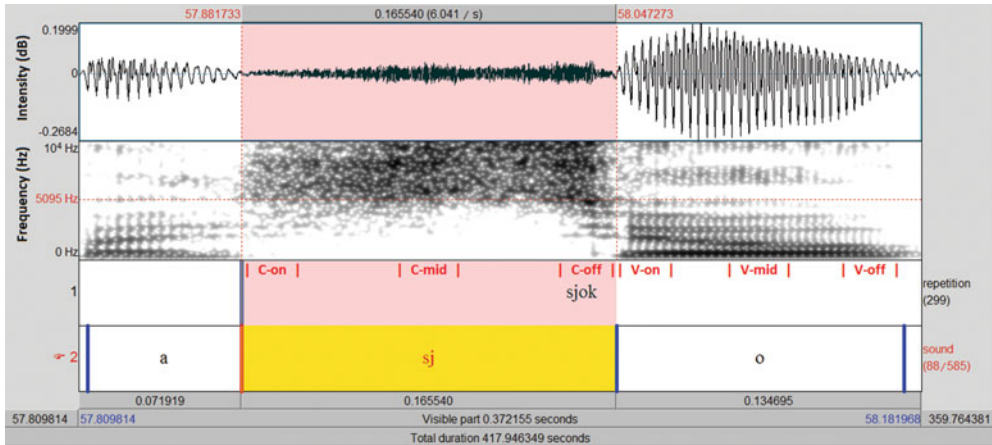
### 4.3 Analysis

The collected audio data were annotated in *Praat* (Boersma & Weenink 2015). All fricatives and following vowels were segmented out based on a visual inspection of the waveform and spectrogram. The onset and the offset of the fricative were taken to be the onset and the offset of fricative noise (appearing on the spectrogram above 2500 Hz), typically coinciding with the lack of vocal fold vibration. The onset and the offset of the vowel before the fricative were

<sup>7</sup> The participants’ native cities represented various regions of the Russian Federation: the Central (Moscow and Moscow oblast), Northwestern (St. Petersburg and Perm), Volga (Kazan, Ulyanovsk, and Volgograd), and Siberian (Novosibirsk) federal districts. The decision not to control for the speakers’ dialect was motivated by the need to identify general acoustic characteristics of Standard Russian fricatives, rather than those that are dialect-specific. It should be noted that that varieties of Standard Russian spoken in larger urban centres are relatively homogenous, typically showing little variation in the phonetic realization of consonants, particularly among younger speakers (Bondarko & Verbitskaya 1987).

<sup>8</sup> The Russian vowels /o/ and /e/ are phonetically lower mid (alternatively transcribed as [ɛ] and [ɔ]), yet higher next to palatalized consonants. The vowel /i/ has a central(ized) allophone [ɨ] after non-palatalized consonants (e.g. Bolla 1981).





**Figure 2** (Colour online) An example of data annotation for a token of the word [sʲok] 'whipped' produced by speaker F1, with approximate analysis windows (see the text) shown below the waveform.

taken to be the onset of vocal fold vibration (the first glottal pulse) and the onset of frication (see above). The onset and the offset of the vowel after the fricative were taken to be the offset of frication and the offset of vocal fold vibration. These measurements were based on the waveform with reference to the corresponding spectrogram, where the boundaries tended to coincide with the onset and offset of the second formant, F2. **Figure 2** presents a sample annotated token of /sʲ/ in [sʲok], indicating the intervals used in the analysis – the frication of /sʲ/ and the following vowel /o/.

Annotated data were subjected to the following measurements using a *Praat* script:

- Consonant duration (in ms).
- Amplitude difference (in dB), calculated by subtracting the amplitude at the midpoint of the fricative (C-mid, see **Figure 2**) from the amplitude taken at the midpoint of the following vowel, or the preceding vowel for word-final fricatives (V-mid), measured using a 25 ms Gaussian window. Lower amplitude difference corresponds to greater intensity of a fricative. The difference, rather than the absolute amplitude, was used to normalize for inter-speaker differences in recording levels (see Jongman, Wayland & Wong 2000, Kochetov & Lobanova 2007).
- Centre of gravity of fricative noise (COG, or the first spectral moment, in Hz), measured at three points in time: at the onset (C-on), the midpoint (C-mid), and the offset of the fricative (C-off), using a 25 ms Gaussian window and a 500 Hz to 10000 Hz pass Hann filter. The windows were either aligned to fricative edges (C-on and C-off) or centred at the midpoint (C-mid). The low cutoff was set to exclude low-frequency room noise or voicing leakage from surrounding vowels, if any (see Zsiga 2000, Nowak 2006). The COG measure was chosen as it is most commonly used to study spectral differences in fricatives (Jongman et al. 2000, Gordon et al. 2002, among others), including Russian fricatives (see **Table 2**; but see Jesus & Shadle 2002 and Spinu, Vogel & Bunnell 2012 on some limitations of and alternatives to the method).
- Formants F1, F2, and F3 (Hz) measured at three points within the following vowel (or the preceding vowel for word-final fricatives) – the onset (V-on), the midpoint (V-mid), and the offset (V-off), using a 25 ms Gaussian window and the Formant (Burg) algorithm. The windows were either aligned to fricative edges (V-on and V-off) or centred at the midpoint (V-mid). Additionally F2–F1 differences were calculated to determine the magnitude of

the difference between palatalized and non-palatalized consonants (Iskarous & Kavitskaya 2010).

The results for Set 1 were analyzed statistically using Repeated Measures Analyses of Variance (RM ANOVAs) separately for each dependent variable (duration, COG, etc.) with within-subjects factors Consonant (/s s<sup>j</sup> ʃ ʃ<sup>j</sup>/) and Position (initial and medial), and between-subjects factor Gender (female, male). These were based on means for each participant, averaged over five vowel contexts (before /a e i o u/). Bonferroni post-hoc tests were performed to investigate differences among the four consonants. Fricative variation due to vowel context was not investigated; however, relevant descriptive statistics by vowel are presented in the appendix (Tables A4–A9). Prior to the RM ANOVA analyses, the results for Set 1 were explored using hierarchical clustering analysis (employing the Ward's cluster method and the measure of squared Euclidean distance). This was done based on mean values for each consonant in initial and medial positions using all dependent variables, with values transformed to *z*-scores. Data from Sets 2 and 3 were analyzed using paired-samples *t*-tests (two-tailed) to examine differences between the same consonants in initial and final positions or before stressed and unstressed vowels. For the reason of space, only significant ( $p \leq .05$ ) effects and interactions are reported.

## 5 Results

### 5.1 Contrasts in prevocalic stressed position

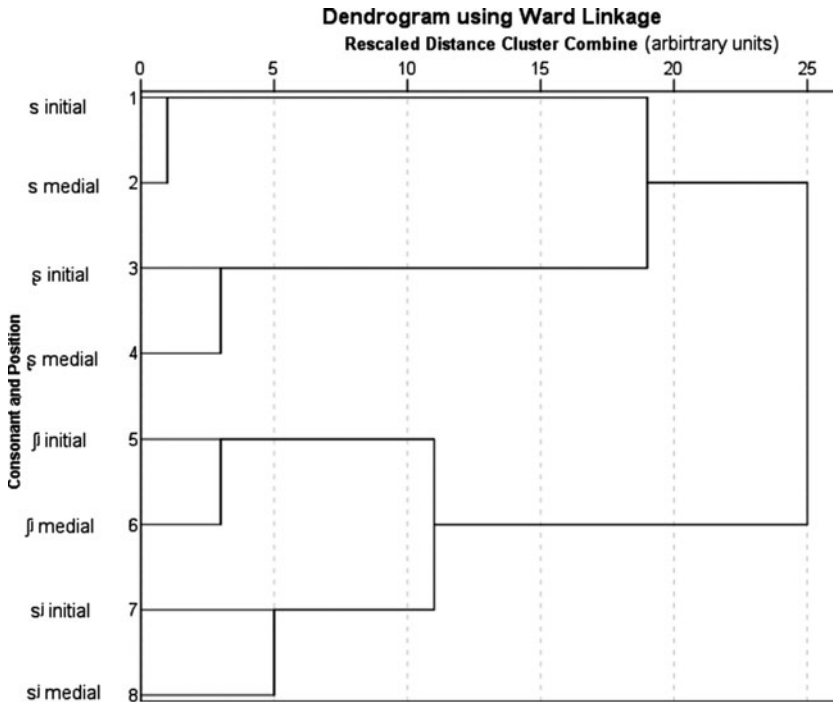
#### 5.1.1 Overview

This section provides an overview of acoustic differences among the four voiceless fricatives, as suggested by the hierarchical clustering analysis. The dendrogram in Figure 3 is based on all acoustic variables used in the analysis, both internal cues (consonant duration, amplitude difference, and COG at three points) and contextual cues (vowel formants F1, F2, and F3 at three points). Distance between cases is shown in the x-axis at the top, increasing from 0 to 25 and above. Higher values indicate more acoustically dissimilar cases, which are represented on the y-axis. Going from right to left, two main clusters emerge – those for non-palatalized and palatalized fricatives. Within each of those, two smaller clusters involve place differences – between /s/ and /s<sup>j</sup>/, and between /ʃ/ and /ʃ<sup>j</sup>/ . Each consonant case is further split into two sub-clusters by position. This clustering analysis thus suggests that the main acoustic difference among the examined phones is secondary articulation, followed by place of articulation, and then by position. Notably, the analysis also suggests that place differences are greater among the non-palatalized fricatives, while positional differences are somewhat greater for the palatalized fricatives.

To what extent is the observed differentiation of the fricatives due to internal or contextual cues as a whole, or due to specific acoustic variables, such as duration, intensity, COG, and formant frequencies? The subsequent sections will explore these questions in detail.

#### 5.1.2 Internal cues: Consonant duration and spectral properties

An RM ANOVA for CONSONANT DURATION showed main effects of Consonant ( $F(3,24) = 25.242, p < .001$ ) and Position ( $F(1,8) = 19.979, p = .002$ ). Bonferroni post-hoc tests revealed that the consonant differences involved /ʃ<sup>j</sup>/ on the one hand and the other fricatives on the other hand (/s/:  $p < .001$ ; /s<sup>j</sup>/:  $p < .01$ ; /ʃ<sup>j</sup>/:  $p < .05$ ). The palatalized posterior fricative was significantly longer than the others, although absolute differences were relatively small, below 20 ms. Means and standard deviations for consonant duration were 143 (28) ms for /s/, 146 (26) ms for /s<sup>j</sup>/, 147 (27) ms for /ʃ/, and 160 (26) ms for /ʃ<sup>j</sup>/ . Based on



**Figure 3** A clustering dendrogram showing the relation among the fricatives /s/, /sʲ/, /ʃ/, and /ʃʲ/ in initial and medial position based on all acoustic cues.

the latter two values, the ratio between the presumed geminate /ʃʲ/ and the singleton /ʃ/ was 1.09. The significant effect of Position was due to the longer duration of initial fricatives compared to intervocalic fricatives. This difference was on average 10 ms. These consonant and position differences are illustrated in Figure 4. Gender was not significant, and there were no significant interactions of any of the factors.

For AMPLITUDE DIFFERENCE, an RM ANOVA showed a main effect of Consonant ( $F(3,24) = 10.135, p < .001$ ). Based on Bonferroni post-hoc tests, significant differences involved /sʲ/ on the one hand and the posterior fricatives /ʃ/ and /ʃʲ/ on the other (both  $ps < .01$ ); the differences between /s/ vs. /ʃ/ and /ʃʲ/ did not reach significance. As seen in Figure 5, anterior fricatives show on average higher amplitude difference (i.e. being less intense) than posterior fricatives. The factors Position and Gender were not significant; there were no significant interactions of any of the factors.

For COG (illustrated in Figure 6), RM ANOVAs were performed at each of the three time points: the onset, the midpoint, and the offset. The analyses revealed a main effect of Consonant at all three points in time ( $F(3,24) = 208.216, p < .001$ ;  $F(3,24) = 200.145, p < .001$ ;  $F(3,24) = 46.885, p < .001$ , respectively), as well as a main effect of Gender at the midpoint and the offset ( $F(1,8) = 7.657, p < .05$ ;  $F(1,8) = 6.823, p = .047$ ). Among the consonants, all pairs were significantly different from each other at the fricative onset: /s/ vs. /ʃ/ and /ʃʲ/ ( $p < .001$ ), /s/ vs. /sʲ/ ( $p < .01$ ), /sʲ/ vs. /ʃ/ and /ʃʲ/ ( $p < .001$ ), /ʃ/ vs. /ʃʲ/ ( $p < .05$ ). COG consonant differences at the midpoint were significant only for the fricative pairs that differed in place: anterior /s/ and /sʲ/ vs. posterior /ʃ/ and /ʃʲ/ (all  $ps < .001$ ). The same was observed for the consonant differences in COG taken at the fricative offset: /s/ vs. /ʃ/ and /ʃʲ/ ( $p < .001$  and  $p < .01$ , respectively), /sʲ/ vs. /ʃ/ and /ʃʲ/ ( $p < .001$  and  $p < .01$ , respectively). Notably, COG differences between /ʃ/ and /ʃʲ/ were not significant at the midpoint and the offset,

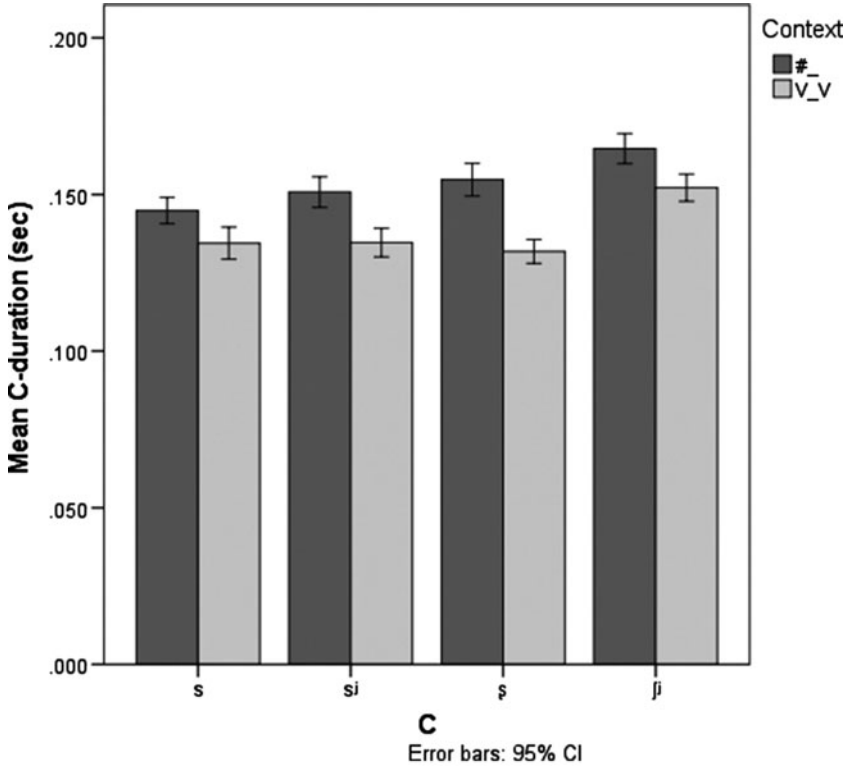


Figure 4 Mean consonant duration (in seconds) values for /s/, /sʲ/, /ʂ/, and /ʃ/ by position (initial and medial).

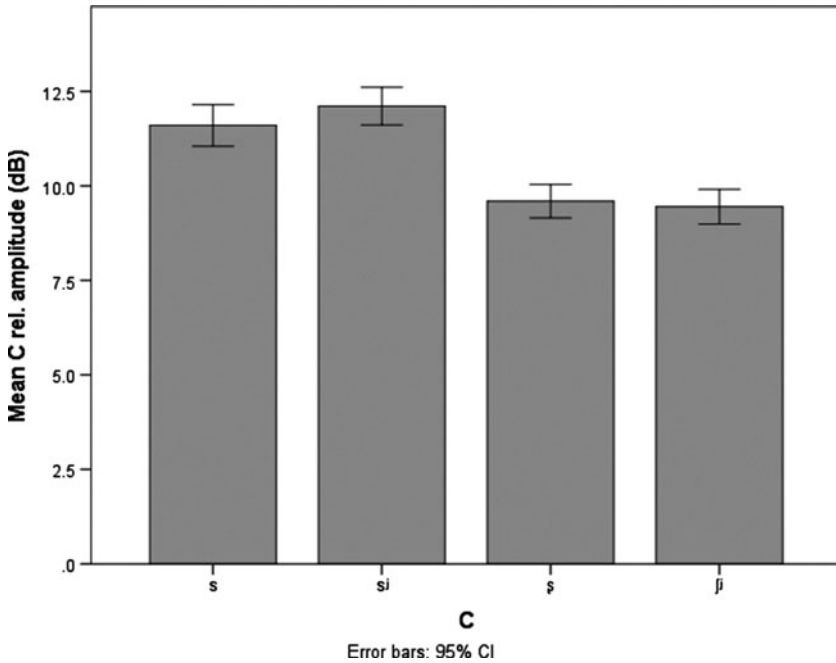
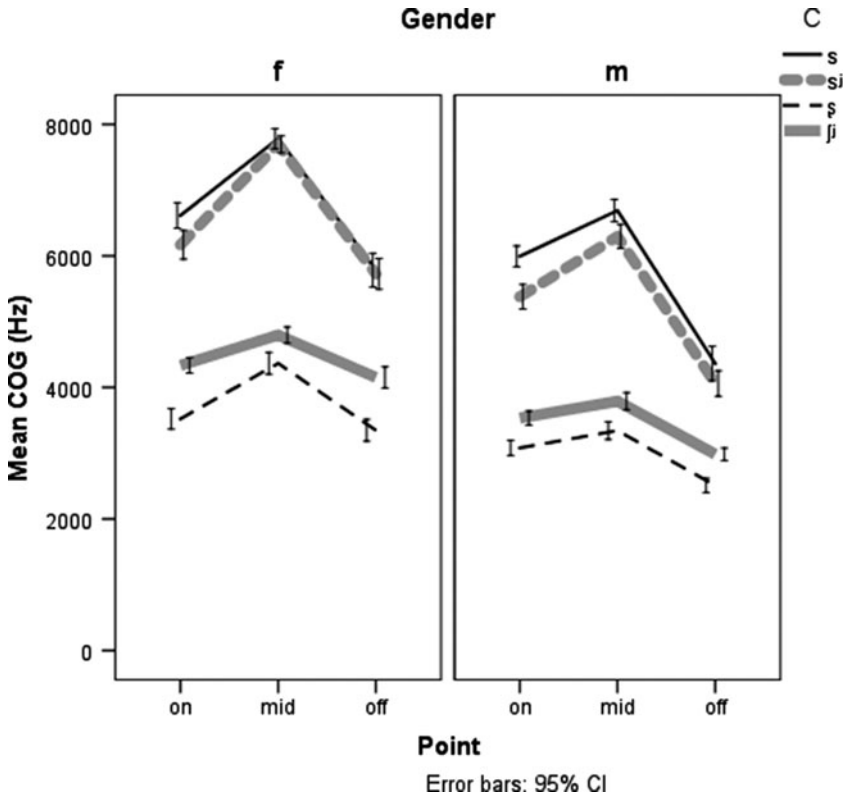


Figure 5 Mean amplitude difference (dB) values for /s/, /sʲ/, /ʂ/, and /ʃ/.



**Figure 6** Mean COG (Hz) values for /s/, /sʲ/, /ʂ/, and /ʃʲ/ at the onset, midpoint, and offset of frication, separately for female and male speakers.

and at best marginally significant at the offset ( $p = .071$ ). On average, COG values at the midpoint were 7236 Hz for /s/, 6989 Hz for /sʲ/, 3884 Hz for /ʂ/, and 4296 Hz for /ʃʲ/. The above-mentioned Gender effect was due to higher COG for the female speakers compared to the male speakers. This difference was on average 1141 Hz at the midpoint and 1284 Hz at the offset. The factor Position was not significant; however, at two time points, there were significant interactions of Consonant  $\times$  Position (offset:  $F(3,24) = 6.169$ ,  $p = .003$ ) and Consonant  $\times$  Position  $\times$  Gender (midpoint:  $F(3,24) = 3.080$ ,  $p = .047$ ). The first interaction was due to lower COG of /s/ in the word-medial position compared to the word-initial position. The second interaction appears to be due to on average higher difference between two posterior fricatives /ʂ/ and /ʃʲ/ (higher COG for the latter) produced by male speakers in the intervocalic position, compared to the word-initial position or to the female speakers' productions.

Figure 6 plots mean COG in time for the four fricatives, separately for female and male speakers. It can be seen that there is a clear distinction between anterior fricatives showing higher COG (above 5000 Hz for females and 4000 Hz for males) and posterior fricatives showing lower COG (below the same thresholds). Within the anterior fricatives, average values are slightly lower for /sʲ/ than /s/ (more so for the males and mainly at the fricative onset). Within the posterior ones, values are on average higher for /ʃʲ/ than /ʂ/. Neither of these differences, however, were significant at the midpoint and the offset. Means and standard deviations for selected internal cue variables by vowel context are given in Tables A4–A6 in the appendix.

**Table 4** A summary of RM ANOVAs and Bonferroni post-hoc results for the factor Consonant: Internal cue variables. '<' and '>' indicate, respectively, higher and lower values; significance: \*\*\* =  $p < .001$ , \*\* =  $p < .01$ , \* =  $p < .05$ , ns = not significant; cells with significant results are shaded.

Variable	Consonant effect	Pairwise comparisons					
		s - $\xi$	s - $j^i$	s - $s^j$	$\xi$ - $j^i$	$\xi$ - $s^j$	$j^i$ - $s^j$
C duration	***	ns	s < $j^i$ ***	ns	$\xi$ < $j^i$ **	ns	$j^i$ > $s^j$ *
Amplitude difference	***	ns	ns	ns	ns	$\xi$ > $s^j$ **	$j^i$ > $s^j$ **
COG on	***	s > $\xi$ ***	s > $j^i$ ***	s > $s^j$ **	$\xi$ < $j^i$ *	$\xi$ < $s^j$ ***	$j^i$ < $s^j$ ***
COG mid	***	s > $\xi$ ***	s > $j^i$ ***	ns	ns	$\xi$ < $s^j$ ***	$j^i$ < $s^j$ ***
COG off	***	s > $\xi$ ***	s > $j^i$ **	ns	ns	$\xi$ < $s^j$ ***	$j^i$ < $s^j$ **

Table 4 provides a summary of statistical results for each of the internal cue dependent variables. It can be seen that the Consonant effect turned out to be highly significant in all analyses. Some pairs of consonants, however, were distinguished by internal cue variables better than others. Specifically, the most acoustically differentiated were the pairs that differ in place ( $/j^i/$  vs.  $/s^j/$  and  $/s/$ ,  $/\xi/$  vs.  $/s^j/$  and  $/s/$ ). Considerably less distinguished were the pairs having the same place and differing in secondary articulation, with  $/s/$  vs.  $/j^i/$  differing in duration and COG at the onset (and marginally in COG at the offset), and  $/s/$  vs.  $/s^j/$  differing only in COG at the onset. Among the variables, COG was useful in distinguishing most contrasts (primarily involving place differences), followed by intensity (place) and duration ( $/j^i/$  vs. the others).

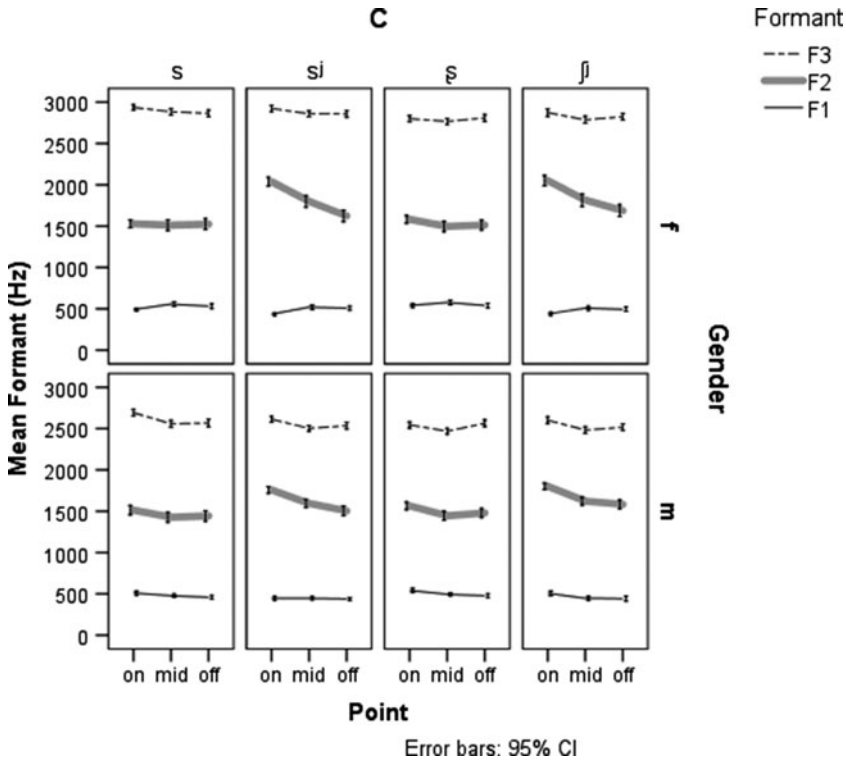
**5.1.3 Contextual cues: Spectral properties of the following vowel**

Turning to contextual cues, RM ANOVAs for VOWEL FORMANTS were performed at each of the three time points: the onset, the midpoint, and the offset (see Figure 7). The results for F1 showed a main effect of Consonant at all three points (onset:  $F(3,24) = 26.586$ ,  $p < .001$ ; midpoint:  $F(3,24) = 25.245$ ,  $p < .001$ ; offset:  $F(3,24) = 5.316$ ,  $p = .006$ ). Significant consonant differences, determined by post-hoc tests, showed lower F1 after palatalized fricatives compared to non-palatalized ones: for all pairs at the onset ( $/s^j/$  vs.  $/s/$ :  $p < .01$ ;  $/s^j/$  vs.  $/\xi/$ :  $p < .001$ ;  $/j^i/$  vs.  $/\xi/$ :  $p < .01$ ) and at the midpoint ( $/s^j/$  vs.  $/s/$  and  $/\xi/$ :  $p < .01$ ;  $/j^i/$  vs.  $/s/$  and  $/\xi/$ :  $p < .001$ ), as well as for one pair at the offset ( $/j^i/$  vs.  $/\xi/$ :  $p < .01$ ). There was a main effect of Gender at the midpoint ( $F(1,8) = 6.311$ ,  $p = .036$ ), with females having higher F1 than males. This difference was on average 70 Hz at the vowel midpoint. The factor Position was not significant, and there were no significant interactions.

RM ANOVAs for F2 showed a main effect of Consonant at all three points (onset:  $F(3,24) = 96.273$ ,  $p < .001$ ; midpoint:  $F(3,24) = 54.918$ ,  $p < .001$ ; offset:  $F(3,24) = 19.065$ ,  $p < .001$ ). Significant consonant differences, determined by post-hoc tests, showed higher F2 for palatalized fricatives compared to the non-palatalized ones: for all pairs at the onset ( $/s^j/$  vs.  $/s/$  and  $/\xi/$ :  $p < .001$ ;  $/j^i/$  vs.  $/s/$  and  $/\xi/$ :  $p < .001$ ) and at the midpoint ( $/j^i/$  vs.  $/s/$ :  $p < .01$ ;  $/j^i/$  vs.  $/\xi/$ :  $p < .001$ ;  $/s^j/$  vs.  $/s/$ :  $p < .01$ ;  $/s^j/$  vs.  $/\xi/$ :  $p < .001$ ), as well as for two pairs at the offset ( $/j^i/$  vs.  $/s/$  and  $/\xi/$ :  $p < .01$ ). There was a main effect of Gender at all three points (onset:  $F(1,8) = 19.174$ ,  $p = .002$ ; midpoint:  $F(1,8) = 17.028$ ,  $p = .003$ ; offset:  $F(1,8) = 7.851$ ,  $p = .023$ ), with female speakers having significantly higher F2 than male speakers. This difference was on average 139 Hz at the vowel midpoint. There were, however, significant Consonant  $\times$  Gender interactions at two points (onset:  $F(3,24) = 11.067$ ,  $p < .001$ ; midpoint:  $F(3,24) = 4.668$ ,  $p < .05$ ). These interactions were due to the greater Gender differences in F2 of the palatalized fricatives. The factor Position was not significant; there were no other significant interactions.

RM ANOVAs for F3 showed a main effect of Consonant at the first two points (onset:  $F(3,24) = 11.456$ ,  $p < .001$ ; midpoint:  $F(3,24) = 13.288$ ,  $p < .001$ ). Post-hoc tests revealed significantly lower F3 at the onset for  $/\xi/$  compared to  $/s/$  and  $/j^i/$  ( $p < .01$  and  $p < .05$ ,





**Figure 7** Mean formant (F1, F2, F3; Hz) values for /s/, /s<sup>ʃ</sup>/, /ʂ/, and /ʃ/ at the onset, midpoint, and offset of the following vowel, separately for female and male speakers.

respectively), and higher F3 at the midpoint for /s/ compared to /ʂ/ and /ʃ/ ( $p < .01$  and  $p < .05$ , respectively). The factor Position was not significant; however, there was a significant Consonant  $\times$  Position interaction at two points (midpoint:  $F(3,24) = 6.540, p = .002$ ; offset:  $F(3,24) = 5.054, p = .007$ ). This interaction was due to the difference between /s<sup>ʃ</sup>/ and /ʃ/ (lower F3 for the latter) word-medially but not word-initially. There was a main effect of Gender at the vowel onset ( $F(1,8) = 10.878, p = .011$ ), with F3 being higher for females than males (by 270 Hz).

To illustrate some of these differences, F1–F3 at three points are plotted in Figure 7, separately by consonant and gender. It can be seen that the main difference between palatalized and non-palatalized consonants is in the F2 transition. This transition starts high for /s<sup>ʃ</sup>/ (on average 1896 Hz) and /ʃ/ (1927 Hz) and gradually declines. The transition for /s/ and /ʂ/ is largely flat (starting at 1520 Hz and 1573 Hz, respectively) or slightly declining. Being quite robust at the onset, F2 differences between palatalized and non-palatalized consonants are still large at the midpoint, and noticeable towards the offset of the vowel. Vowels after palatalized consonants also have somewhat lower F1, although these differences are much smaller in magnitude (with values being on average 438 Hz for /s<sup>ʃ</sup>/, 473 Hz for /ʃ/ vs. 500 Hz for /s/ and 532 Hz for /ʂ/). Altogether, palatalized consonants exhibit a higher separation of F1 and F2 (1458 Hz and 1454 Hz) compared to non-palatalized consonants (1020 Hz and 1041 Hz). F3 is slightly lower for /ʂ/ than for the other fricatives (on average 2671 Hz for /ʂ/ vs. 2815 Hz for /s/, 2765 Hz for /s<sup>ʃ</sup>/, and 2737 Hz for /ʃ/).

Table 5 provides a summary of statistical results for each of the contextual dependent variables. It can be seen that the Consonant effect was significant in all analyses involving F1 and F2, as well as in all but one analysis of F3. Vowel duration was not affected by

**Table 5** A summary of RM ANOVAs and Bonferroni post-hoc results for the factor Consonant: Contextual cue variables. '<' and '>' indicate, respectively, higher and lower values; significance: \*\*\* =  $p < .001$ , \*\* =  $p < .01$ , \* =  $p < .05$ , ns = not significant; cells with significant results are shaded.

Variable	Consonant effect	Pairwise comparisons					
		s - $\xi$	s - $j^i$	s - $s^j$	$\xi$ - $j^i$	$\xi$ - $s^j$	$j^i$ - $s^j$
F1 on	***	ns	ns	s > $s^j$ **	$\xi$ > $j^i$ **	$\xi$ > $s^j$ ***	ns
F1 mid	***	ns	s > $j^i$ ***	s > $s^j$ **	$\xi$ < $j^i$ ***	$\xi$ > $s^j$ **	ns
F1 off	**	ns	ns	ns	$\xi$ < $j^i$ *	ns	ns
F2 on	***	ns	s > $j^i$ ***	s > $s^j$ ***	$\xi$ < $j^i$ ***	$\xi$ > $s^j$ ***	ns
F2 mid	***	ns	s > $j^i$ **	s > $s^j$ **	$\xi$ < $j^i$ ***	$\xi$ > $s^j$ ***	ns
F2 off	***	ns	s > $j^i$ *	ns	$\xi$ < $j^i$ **	ns	ns
F3 on	***	s > $\xi$ **	ns	ns	$\xi$ < $j^i$ *	ns	ns
F3 mid	***	s > $\xi$ **	s > $j^i$ *	ns	ns	ns	ns
F3 off	ns	ns	ns	ns	ns	ns	ns

consonant contrasts. Considering significant differences among the consonants, the pairs that differed in palatalization were the most distinct: / $j^i$ / vs. / $\xi$ / and / $s$ /, and / $s^j$ / vs. / $\xi$ / and / $s$ / (in declining order). Place-only differences were either distinguished by F3, as for / $s$ / vs. / $\xi$ /, or not distinguished at all, as for / $j^i$ / vs. / $s^j$ /. Means and standard deviations for selected contextual cue variables by vowel context are given in Tables A7–A9 in the appendix.

### 5.1.4 Summary

To summarize, the analyses of internal cue variables showed a robust differentiation of fricatives by place (in COG and amplitude difference). Contrasts within places of articulation (anterior / $s$ / vs. / $s^j$ / and posterior / $\xi$ / vs. / $j^i$ /) were not significantly distinguished by internal cue variables, with the exception of COG at the onset (both pairs) and duration (the latter pair). On the other hand, the analyses of contextual cue variables showed a robust differentiation of secondary articulation differences – / $s$ / vs. / $s^j$ / and / $\xi$ / vs. / $j^i$ / (in both F1 and F2), but a rather limited differentiation of the primary place (only / $s$ / vs. / $\xi$ / in F3).

Differences in position involved primarily consonant duration (initial > medial), as well as COG or formants for specific consonants. Some of these differences can be attributed to positional enhancement (e.g. higher COG in the more prominent word-initial position), while others could reflect contextual differences specific to lexical items (e.g. presence of postvocalic palatalized consonants). Gender differences were found for fricative COG and formants F1–F3, with females showing higher values than males for both spectral properties. These differences are fully expected, given the well-established physiological and sociophonetic differences in these parameters (e.g. Jongman et al. 2000, Munson et al. 2006). Somewhat unexpectedly, however, these differences were greater for some consonants than others – specifically for anterior fricatives in COG and for palatalized fricatives in F1 and F2. This resulted in the place and secondary articulation contrasts being more dispersed for females than males.

Overall, COG and the first two formants emerged as the primary acoustic characteristics of the four-way contrast, with the latter measure being crucial to distinguishing the primary place of articulation (anterior vs. posterior), and the former measure distinguishing the secondary articulation (palatalized vs. non-palatalized) of the fricatives. As shown in Figure 8, these two parameters clearly define the contrasts in both word-initial and word-medial positions. The same can be said about individual results shown in Figure 9 (averaged by position): all 10 speakers show a relatively symmetrical dispersion of the fricative contrasts, although to a lesser extent in the F2–F1 dimension for some of the male speakers. (See Figure A1 in the appendix for the results by vowel context.)

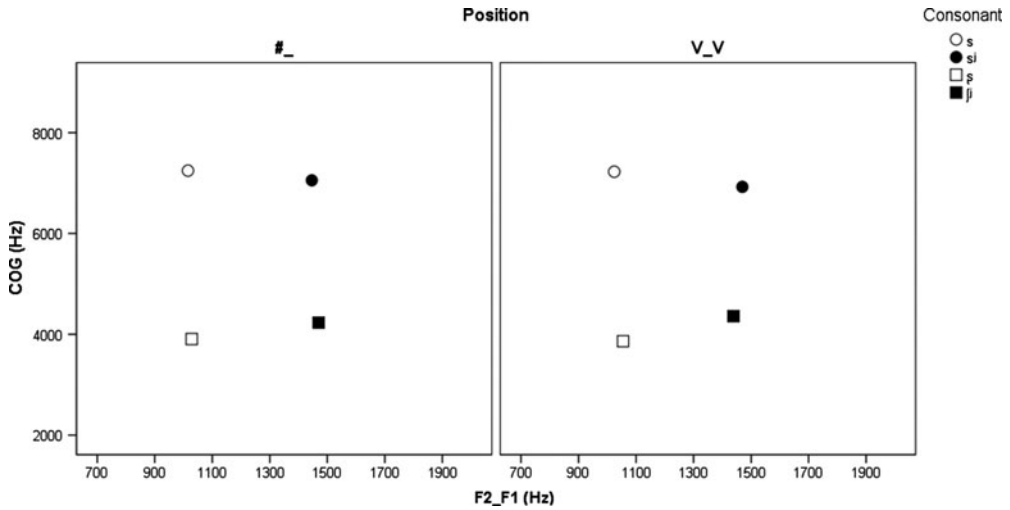


Figure 8 Plots showing mean COG at the midpoint (Hz) by mean F2-F1 difference at the vowel onset (Hz) for the fricatives /s/, /sʲ/, /s̺/, and /ʃ/ separately by position – word-initial and word-medial (averaged for vowel contexts and speakers).

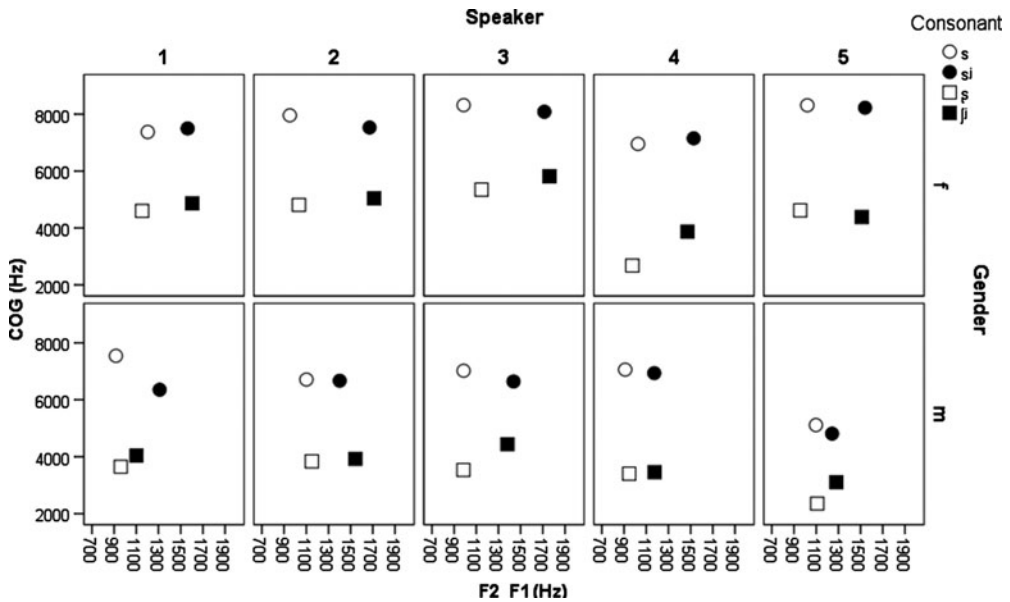


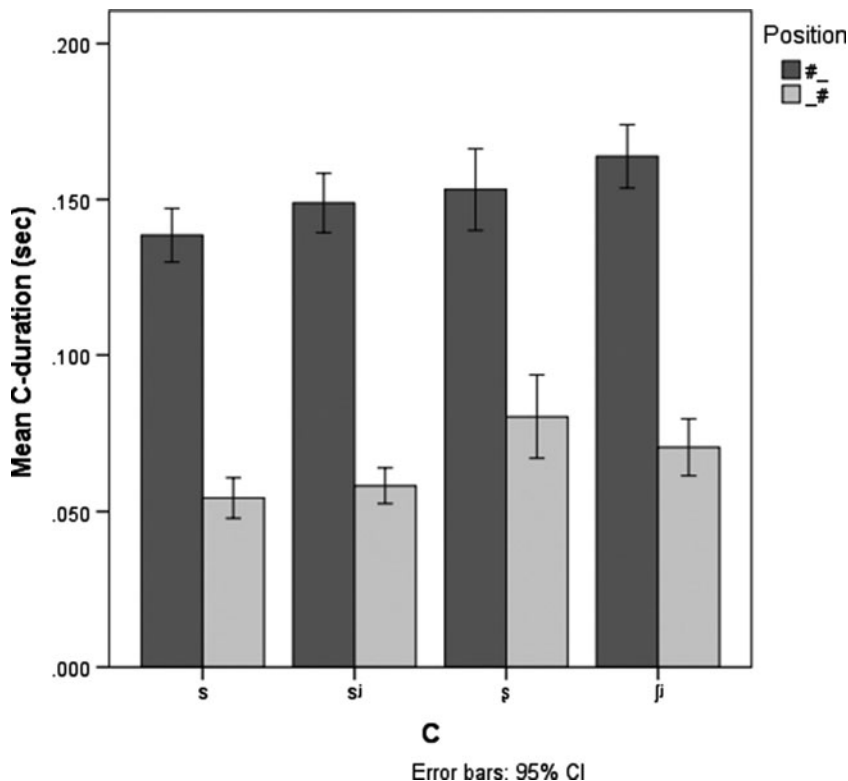
Figure 9 Plots showing mean COG at the midpoint (Hz) by mean F2-F1 difference at the vowel onset (Hz) for the fricatives /s/, /sʲ/, /s̺/, and /ʃ/ separately for each speaker (averaged for positions and vowel contexts).

## 5.2 Fricatives in other contexts

This section presents results for the remaining comparisons – initial/final position and stress.

### 5.2.1 Word-final position

Words from Set 2 (see appendix Table A2) were investigated using paired samples *t*-tests (t-tailed) to determine differences between word-initial and word-final fricatives. Compared to initial fricatives, the same consonants in final position had shorter duration

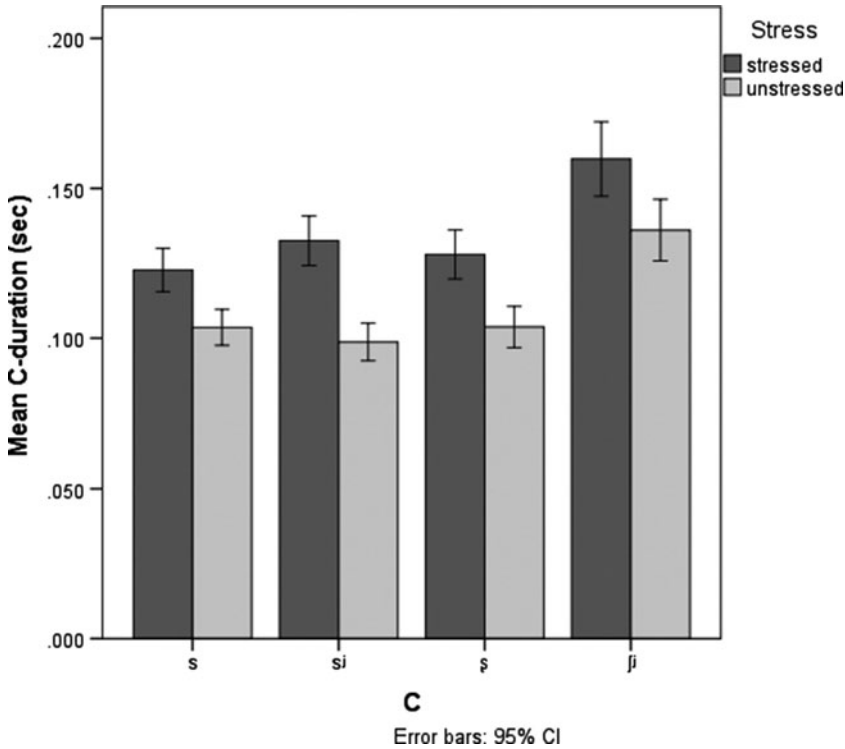


**Figure 10** Mean consonant duration (in seconds) values for /s/, /sʲ/, /ʃ/, and /ʃʲ/ by position, initial and final (both next to vowel /o/).

(/s/:  $t(9) = -16.680, p < .001$ ; /sʲ/:  $t(9) = -11.673, p < .001$ ; /ʃ/:  $t(9) = -9.197, p < .001$ ; /ʃʲ/:  $t(9) = -17.195, p < .001$ ). As can be seen in [Figure 10](#), this reduction was quite substantial, ranging between 73 ms and 91 ms (i.e. 50–60% of the initial duration of the consonants), being the largest for /ʃʲ/. This resulted in the /ʃʲ/ duration being shorter than for its ‘short’ non-palatalized counterpart /ʃ/ (73 ms vs. 80 ms), effectively neutralizing the length contrast.<sup>9</sup> COG at the midpoint of /sʲ/ was lower in final position compared to initial position ( $t(9) = 5.386, p < .001$ ). The difference was about 950 Hz.

All word-final fricatives showed higher F2 (measured at the vowel offset) compared to word-initial fricatives (measured at the vowel onset) (/s/:  $t(9) = 4.486, p < .01$ ; /sʲ/:  $t(9) = 3.192, p < .05$ ; /ʃ/:  $t(9) = 4.896, p < .01$ ; /ʃʲ/:  $t(9) = 2.939, p < .05$ ). This difference was on average 170 Hz, and is possibly due to differences in the make-up of word-initial and word-final stimuli (with the other vowel-adjacent consonant being a velar or alveolar/labial, respectively). In addition, word-final anterior fricatives showed lower F3 than their counterparts in word-initial position (/s/:  $t(9) = -3.051, p < .05$ ; /sʲ/:  $t(9) = -7.773, p < .001$ ), by on average 240 Hz. These differences appear to be due to F3 lowering by the preceding /r/ for the items [zbros] and [brosʲ].

<sup>9</sup> An examination of individual duration results showed that the final /ʃʲ/ was on average shorter than the final /ʃ/ for all but two speakers. Even for the latter two speakers, the /ʃʲ/ > /ʃ/ difference was rather small, 8–10 ms.



**Figure 11** Mean consonant duration (in seconds) values for /s/, /s<sup>j</sup>/, /ʃ/, and /jʃ/ in two stress conditions (stressed and unstressed).

### 5.2.2 Unstressed vowel context

Words from Set 3 (see Table A3) were examined for the effect of stress on the fricatives. Compared to their stressed counterparts, fricatives in unstressed syllables were characterized by shorter duration (/s/:  $t(1,9) = 4.935, p < .01$ ; /s<sup>j</sup>/:  $t(9) = 8.870, p < .001$ ; /ʃ/:  $t(9) = 7.815, p < .001$ ; /jʃ/:  $t(9) = 3.677, p < .01$ ) and lower F1 at the onset of the vowel (/s/:  $t(9) = 4.935, p < .01$ ; /s<sup>j</sup>/:  $t(9) = 5.286, p < .01$ ; /ʃ/:  $t(9) = 4.006, p < .01$ ; /jʃ/:  $t(9) = 3.677, p < .01$ ). As seen in Figure 11, the duration differences were relatively small, on average 25 ms, which corresponds to a 20% reduction in duration in unstressed syllables. This reduction did not affect the duration difference between /jʃ/ and the other fricatives, as the former was still on average longer than the other fricatives (136 ms vs. 99–104 ms). It should be noted that the /ʃ/ vs. /jʃ/ duration ratio observed here (1.25–1.30) is somewhat higher than in the analysis reported in Section 5.1.2 above. This difference possibly reflects some variation in the duration of /jʃ/ by position and vowel context. The above-mentioned F1 differences were on the magnitude of 100 Hz, and seem to reflect some raising (centralization) of /a/ to [ʌ] or [ə] in unstressed syllables. In addition, vowels adjacent to palatalized fricatives in unstressed position showed lower F3 than vowels adjacent to the same consonants in stressed position (/s<sup>j</sup>/:  $t(9) = 3.465, p < .01$ ; /jʃ/:  $t(9) = 3.464, p < .01$ ). This difference, the source of which is unclear, was about 140 Hz.

### 5.2.3 Summary

The results presented in this section showed that duration was important in signaling positional and stress effects, with word-final and unstressed fricatives being considerably shorter than their word-initial and stressed counterparts. An important result of this positional

shortening is the neutralization of the durational distinction between /ʃ<sup>j</sup>/ and the other fricatives, the contrast that was already relatively small in the more prominent initial and medial positions. Other positional differences involved less consistent variation in COG, F2, and F3, possibly reflecting effects of non-adjacent consonants or vowel centralization. The positional shortening of final consonants seems unexpected given the literature documenting the propensity of the former to lengthening (e.g. Beckman & Edwards 1990). It should be noted, however, that lengthening usually involves utterance-final consonants, while the word-final consonants in this study were utterance-medial, embedded in a carrier phrase. Shortening of word-final fricatives was also reported by Kochetov & Lobanova (2007) for Komi-Permyak.

## 6 Discussion and conclusion

The results of the study demonstrate that Russian voiceless sibilant fricatives are robustly distinguished by spectral differences – the frequency concentration of the fricative noise and the formants of an adjacent vowel. The higher COG for anterior fricatives /s<sup>j</sup>/ and the lower COG for posterior fricatives /ʃ<sup>j</sup>/ is consistent with previous analyses of the Russian contrast (Bolla 1981, Funatsu & Kiritani 1998, Zsiga 2000, Padgett & Žygis 2007, Kochetov & Radišić 2009), as well as is expected based on what we know about the articulation of these consonants (e.g. Skalozub 1963, Bolla 1981, Avanesov 1984, Kedrova et al. 2008). What is new, however, is the finding that the anterior/posterior differences are substantial and significant throughout the fricative (and regardless of the vowel context). While this is not surprising for the non-palatalized fricatives /s/ and /ʃ/ (see Gordon et al. 2002 on place differences in sibilants across languages), it is notable for the palatalized pair, /s<sup>j</sup>/ and /ʃ<sup>j</sup>/. Cross-linguistically, anterior coronals are prone to shifting to the posterior place when overlapped by or adjacent to a palatal articulation (Bhat 1978, among others). We could therefore expect /s<sup>j</sup>/ to spectrally approach /ʃ<sup>j</sup>/, at least its offset. This is clearly not the case in our data: the COG of /s<sup>j</sup>/ was consistently high (/s/-like) throughout the frication, including the fricative offset. This is in line with Zsiga's (2000) results, which also showed a temporally stable COG for Russian /s<sup>j</sup>/, in contrast with a gradually decreasing COG in English /s+j/ sequences (e.g. /sj/ → [sʃ] in *miss you*). The lack of a similar palatal assimilation in Russian presumably reflects a language-particular constraint on the coarticulation of the overlapping tongue tip and tongue body gestures, effectively suppressing the otherwise expected categorical or gradient change in place. As palatalized consonants in Russian also contrast with consonant + glide sequences (e.g. [s<sup>j</sup>eɪ] 'sat down' vs. [sʃeɪ] 'has eaten'), it would be interesting to examine how these contrasts are implemented acoustically and articulatorily.

Also new is the finding of within-category differences in COG at least at the onset of frication: the palatalized /ʃ<sup>j</sup>/ had higher COG than the non-palatalized /ʃ/, while the palatalized /s<sup>j</sup>/ had lower COG than the non-palatalized /s/. These differences (and particularly for /ʃ/ vs. /ʃ<sup>j</sup>/) were also observed numerically at the midpoint and the offset of frication, although they did not reach significance. A non-significant tendency towards this was previously noted in Padgett & Žygis (2007) and observed in Kochetov & Radišić (2009), the studies examining sibilant fricatives in the context before /a/.

Another internal acoustic property – fricative noise intensity (reflected in the measure of amplitude difference) – was also found to contribute to the realization of primary place contrasts. This is in part consistent with Kochetov & Radišić's (2009) findings (where both posterior fricatives before /a/ were more intense) and different from Bolla's (1981) observation (where palatalized fricatives were louder; not controlled for the following vowel).

While fricative noise spectral differences were crucial to distinguishing place contrasts, formant differences during the vowel were important for the palatalized vs. non-palatalized



distinction. Higher F2 of an adjacent (and particularly the following) vowel is a well-known correlate of consonant palatalization, reflecting the decreasing size of the front cavity (Fant 1960, Shupljakov et al. 1968, Derkach et al. 1970, Purcell 1979, among others). To a lesser degree this applies to lower F1, corresponding to the raising of the tongue (e.g. Derkach et al. 1970). The resulting F2–F1 difference at the vowel onset in our data was found to be consistently higher for palatalized consonants than non-palatalized ones, being on average 425 Hz (and ranging from 345 Hz to 510 Hz depending on the vowel context). This is comparable for the difference obtained by Purcell (1979) for Russian coronal stops across five vowel contexts (490 Hz) and by Kochetov & Radišić (2009) for the fricatives in the context of /a/ (550 Hz). Moreover, the current results showed that formant differences were not limited to the onset of the vowel, but were almost as robust at vowel midpoint, and – at least for /j/ – still significant at the vowel offset. This reflects the large temporal extent of the palatalization gesture in Russian, as well as a seemingly greater degree of palatalization of /j/ compared to /sʲ/. The lower F3 for /s/ (compared to /sʲ/) can be attributed to the consonant tip-up (retroflex-like) articulation. This difference, however, was fairly small (e.g. 200 Hz for for the /s/ – /sʲ/ pair), in comparison to the robust differences found for the true retroflex/non-retroflex contrasts as in Toda (Gordon et al. 2002). This suggests that the Russian /s/ is relatively weakly retroflexed, comparable to apical post-alveolar fricatives in other Slavic and Finno-Ugric languages (Hamann 2004, Kochetov & Lobanova 2007). Given this, an alternative IPA transcription for the Russian non-palatalized posterior fricative would be /ʃ/ (apical post-alveolar; see Lee & Zee 2003 for Standard Chinese).

Overall, the study shows that the primary acoustic distinction among the four Russian consonants is that of the secondary articulation, followed by primary place, and then positional differences (see the results of the hierarchical clustering analysis in Figure 3). Interestingly, the results also show that place differences are greater among the non-palatalized fricatives, while positional differences are somewhat greater for the palatalized fricatives. This possibly reflects the lesser stability and greater variability of palatalized consonants in the system of contrasts. The finding that the palatalized /sʲ/ and /j/ are distinguished from their non-palatalized counterparts /s/ and /ʃ/ primarily by F1 and F2 of adjacent vowels is important. It explains the contextual restrictions on the distribution of palatalized fricatives. Recall from Table 3 that /sʲ/ and /j/ are partially restricted word-medially before consonants (V\_C; but not before vowels, C\_V) and partially or fully avoided word-finally after consonants (C\_#) and word-initially before consonants (#\_C). Note that the last three contexts lack the following vowels, and thus the cues to the palatalization distinction are considerably reduced (to the preceding vowel or sonorant consonant transitions in V\_C and C\_#) or almost absent (in #\_C). Overall, this provides evidence for the role of acoustic cues in the distribution of phonological contrasts (Steriade 1997; see Kochetov 2006 on Russian stops). Further, while phonotactically, Russian /s/ and /j/ pattern just like other non-palatalized or palatalized consonants, they behave differently in alternations (Timberlake 2004; see Section 1 above). The ‘immutable’ status of posterior fricatives can be attributed to their somewhat greater acoustic contrast compared to /s/ vs. /sʲ/, as manifested in on average larger COG differences and a longer span of F1/F2 differences across the following vowel.

The overall robust and symmetrical differentiation of the Russian sibilant contrasts across the examined contexts provides support for the acoustic/perceptual dispersion approach to fricative inventories (Padgett & Žygis 2007 on Russian and Polish; Žygis & Padgett 2010 on Polish). Recall that Russian fricatives have evolved to become sufficiently distinct from each other (see Figure 1). As our results show, this enhancement must have proceeded primarily along the dimensions of F1/F2 vowel formants and frequency of fricative noise (although somewhat differently than in other Slavic languages; see Padgett & Žygis 2007 on Polish). This study thus further extends the application of the concepts of perceptual optimization and dispersion (Liljencrants & Lindblom 1972) to consonants, beyond the more familiar

cases of vowel inventories (see Lindblom & Maddieson 1988, Flemming 2002; but see Ohala 1979).

Thus it appears that the Russian sibilant fricative inventory has evolved, in Lindblom & Maddieson's (1988) words, 'to achieve maximal perceptual distinctiveness at minimum articulatory cost'.

Somewhat unexpectedly, duration turned out to contribute rather little to distinguishing the fricative contrasts. Previous descriptive phonetic accounts have often considered /ʃ/ to be geminate (i.e. /ʃː/), being considerably longer than the other fricatives. Bolla's (1981) measurements of voiceless fricatives (not controlled for vowel context) showed this to be true, with /ʃ/ being 1.55 times longer than /s/. A more recent study by Kochetov & Radišić (2009), however, found smaller differences, with /ʃ/ (before /a/) being 1.23 times longer than /s/, but not significantly different from /sʲ/. It was not clear to what extent those results were reflecting dialect or vowel context differences. Our current results, nevertheless, are fairly consistent with the latter study, showing even smaller, albeit significant, durational differences, with the average ratio of 1.09 for /ʃ/ vs. /s/. The analysis of a smaller set of data in Section 5.2.2 above showed a somewhat higher ratio of 1.25–1.30, likely reflecting some variation in the realization of /ʃ/ across positions and vowel contexts. Overall, however, these ratios are considerably lower than ratios reported for Russian contrastive and morphologically-conditioned geminates by Dmitrieva (2017), which are in the 1.6–1.8 range. This suggests that /ʃ/ in contemporary Russian is only marginally distinguished by duration, which is in contrast to the many mid-to-late 20th century descriptive and phonetic accounts (Jones & Ward 1969, Matusevich 1976, Bolla 1981, Avanesov 1984). The marginal utility of duration is further confirmed by positional comparisons, showing that the word-final /ʃ/ was not different in duration from the other fricatives. Previous descriptive accounts mentioned the optionality of the /ʃ/ degemination in this context (Timberlake 2004). Our results, however, suggest that the process is no longer optional, as the lack of contrast was exhibited both at the level of the group and individual speakers. Overall, the results suggest a change in progress, likely reflecting a fuller integration of /ʃ/ in the palatalized vs. non-palatalized system of contrasts, where consonant duration does not play a role. This in fact may be another example of the application of the principles of minimizing articulatory cost (Lindblom & Maddieson 1988) or enforcing feature economy (Clements 2003). At the same time, consonant duration in our results was useful in distinguishing positional variation, with considerably shorter values exhibited by all consonants in word-final and unstressed positions, compared to initial (stressed) position. Similar effects of word-final position on fricative duration were previously reported for Komi-Permyak (Kochetov & Lobanova 2007), while stress effects on stop duration were documented for Russian (Zsiga 2000).

To conclude, this study provides an acoustic snapshot of the complex and historically dynamic set of voiceless sibilant fricative contrast in contemporary Russian. Further phonetic work on complex sibilant inventories across languages is expected to provide new insights in the mechanisms underlying their historical development and to further enrich phonetic typology of fricative contrasts.

## Acknowledgments

The paper has benefited greatly from insightful comments and suggestions by Shigeto Kawahara, Mayuki Matsui, and anonymous Journal of International Phonetic Association reviewers. Thanks to Milica Radišić for assistance with the collection, annotation, and extraction of data. The research was supported by grants from the Social Sciences and Humanities Research Council of Canada (#416-2006-1006 and #435-2015-2013).

## Appendix. Materials and means by vowel context

**Table A1** Set 1 of the stimuli: words with four fricatives in word-initial and word-medial position before five stressed vowels.

Vowel	Position	Consonant			
		s	sʲ	ɕ	ʃ
a	#_	'samʲi сами '(them)selves'	'sʲadʲit сядет '(he) will sit'	'ɕalʲi шали 'shawls'	'ʃ'avʲiʲiʲ щавель 'sorrel'
	V_V	za'sada засада 'ambush'	za'sʲadu засяду 'I will sit down'	mʲʂata мышата 'baby mice'	pa'ʃʲada пощада 'mercy'
e	#_	ser сэр 'sir'	sʲer сер 'grey'	ɕest шест 'pole'	ʃʲedr щедр 'generous'
	V_V	at e'sera от эсэра 'from an SR'	padza'sʲela подзасела '(she) got caught'	aku'ɕeri акушеры 'obstetricians'	is pʲi'ʃʲeri из пещеры 'from a cave'
i	#_	sʲit сыт 'full, satiated'	sʲil сил 'forces, GEN'	ɕit шит 'sewn'	ʃʲit щит 'shield'
	V_V	kra'si красы 'beauty, GEN'	pra'sʲi проси 'ask, IMP'	pʲi'ʲɕi пляши 'dance, IMP'	pla'ʃʲi плащи 'raincoats'
o	#_	sok Сок 'juice'	sʲok сёк 'whipped'	ɕok шок 'shock'	ʃʲok щёк 'cheeks, GEN'
	V_V	unʲi'son унисон 'unison'	unʲi'sʲon унесён 'carried away'	razrʲi'ʃon разрешён 'allowed'	razmʲi'ʃʲon размещён 'placed'
u	#_	'sutkʲi сутки 'day and night'	'sʲutkʲin Сюткин 'Syutkin (name)'	'ɕutkʲi шутки 'jokes'	'ʃʲukʲi щуки 'pikes'
	V_V	v na'su в носу 'in the nose'	la'sʲu лосю 'moose, DAT'	na'ɕu ношу 'I wear'	ta'ʃʲu тащу 'I drag'

**Table A2** Set 2 of the stimuli: words with four fricatives in word-initial and word-final position.

Vowel	Position	Consonant			
		s	s <sup>j</sup>	ʂ	ʃ <sup>j</sup>
o	#_	sok	s <sup>j</sup> ok	ʂok	ʃ <sup>j</sup> ok
		сок 'juice'	сѣк 'whipped'	шок 'shock'	щѣк 'cheeks, GEN'
	_#	zbros	bros <sup>j</sup>	broʂ	xvoʃ <sup>j</sup>
		сброс 'reset'	брось 'drop it'	брошь 'brooch'	хвоц 'equisetum'

**Table A3** Set 3 of the stimuli: words with four fricatives before stressed and unstressed vowels. Vstr = stressed vowel, Vunstr = unstressed vowel.

Vowel	Position	Consonant			
		s	s <sup>j</sup>	ʂ	ʃ <sup>j</sup>
a	V_Vstr	za'sada	za's <sup>j</sup> adu	m <sup>i</sup> 'ʂata	pa'ʃ <sup>j</sup> ada
		засада 'ambush'	засяду 'I will sit down'	мышата 'baby mice'	поцада 'mercy'
	V_Vunstr	'fkusa	'gus <sup>j</sup> a	'gruʂa	'guʃ <sup>j</sup> a
		вкуса 'taste, GEN'	гуся 'goose, GEN'	груша 'pear'	гуща 'thick, n'

**Table A4** Mean consonant duration (ms) for four consonants, by vowel context and gender, averaged over positions and speakers.

Consonant	Gender	_a		_e		_i		_o		_u	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
s	f	135	27	164	45	157	29	140	28	141	29
s <sup>j</sup>	f	139	25	155	31	156	22	147	30	155	33
ʂ	f	145	26	153	25	166	30	150	34	147	30
ʃ <sup>j</sup>	f	159	35	169	30	168	31	155	29	161	20
s	m	128	17	146	19	148	24	133	25	136	21
s <sup>j</sup>	m	135	19	156	24	152	26	130	12	138	22
ʂ	m	136	19	138	17	157	19	140	26	134	17
ʃ <sup>j</sup>	m	166	29	158	17	166	23	151	20	148	17

**Table A5** Mean amplitude difference (dB) for four consonants, by vowel context and gender, averaged over positions and speakers.

Consonant	Gender	_a		_e		_i		_o		_u	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
s	f	11.0	6.0	10.2	5.7	9.7	6.6	13.0	5.6	10.3	5.1
s <sup>j</sup>	f	12.5	6.0	11.3	5.3	9.5	4.7	13.2	4.2	10.0	5.4
ʂ	f	8.7	4.3	9.3	4.1	6.8	3.9	10.5	4.6	8.9	4.3
ʃ <sup>j</sup>	f	8.5	3.8	8.4	4.4	7.2	4.4	9.2	4.5	7.2	3.5
s	m	12.3	5.3	11.8	4.4	11.7	4.3	13.8	3.9	12.5	3.6
s <sup>j</sup>	m	14.3	3.2	13.6	3.9	11.6	3.6	14.6	2.9	12.8	3.4
ʂ	m	10.7	4.0	10.8	4.2	9.9	3.6	11.4	3.5	10.5	4.7
ʃ <sup>j</sup>	m	12.0	3.0	13.1	4.0	10.1	3.0	11.3	3.0	9.7	4.2

**Table A6** Mean COG (Hz) at the frication midpoint for four consonants, by vowel context and gender, averaged over positions and speakers.

Consonant	Gender	_a		_e		_i		_o		_u	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
s	f	8091	344	8167	530	8152	455	7181	783	7319	1329
s <sup>j</sup>	f	7930	447	8037	322	8012	310	7370	670	7133	958
ʂ	f	4076	1000	4428	916	4207	970	4659	1056	4690	1304
ʃ <sup>j</sup>	f	4826	674	4845	694	4848	711	4742	938	4717	777
s	m	6693	503	7026	1116	6776	897	6477	1367	6475	1218
s <sup>j</sup>	m	6112	995	6688	905	6572	957	6459	1202	5580	1360
ʂ	m	3207	513	3181	609	3390	450	3713	1360	3287	682
ʃ <sup>j</sup>	m	4126	1031	3882	616	3754	491	3578	574	3641	513

**Table A7** Mean F1 (Hz) at the vowel onset for four consonants, by vowel context and gender, averaged over positions and speakers.

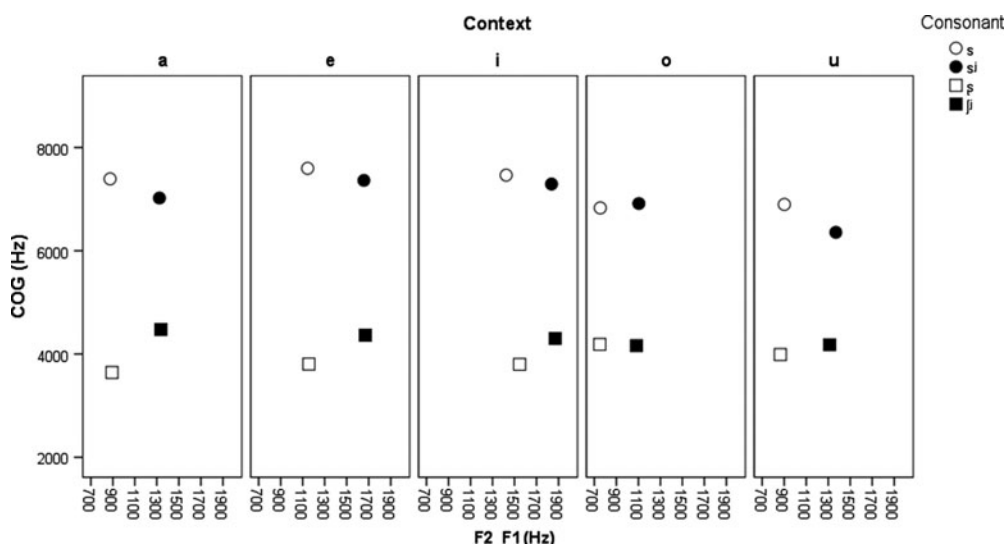
Consonant	Gender	_a		_e		_i		_o		_u	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
s	f	642	32	510	27	401	33	495	47	405	42
s <sup>j</sup>	f	552	43	410	31	358	54	481	48	370	62
ʂ	f	644	83	577	62	429	96	554	35	437	60
ʃ <sup>j</sup>	f	527	64	420	49	388	97	482	44	377	44
s	m	594	74	487	61	402	105	555	58	508	166
s <sup>j</sup>	m	489	47	438	96	382	123	494	77	409	97
ʂ	m	635	49	480	54	408	64	664	93	486	124
ʃ <sup>j</sup>	m	572	73	456	107	455	154	570	104	482	134

**Table A8** Mean F2 (Hz) at the vowel onset for four consonants, by vowel context and gender, averaged over positions and speakers.

Consonant	Gender	_a		_e		_i		_o		_u	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
s	f	1599	84	1746	86	1835	113	1256	137	1205	170
s <sup>j</sup>	f	2041	71	2318	117	2419	138	1662	57	1746	159
ʂ	f	1578	105	1788	103	1975	155	1324	87	1241	133
ʃ <sup>j</sup>	f	2044	201	2370	174	2480	147	1628	134	1732	211
s	m	1387	122	1542	110	1815	155	1301	240	1511	308
s <sup>j</sup>	m	1647	80	1842	137	1993	110	1521	168	1771	76
ʂ	m	1486	140	1582	100	1953	103	1392	241	1411	193
ʃ <sup>j</sup>	m	1728	134	1845	106	2103	62	1591	230	1752	131

**Table A9** Mean F3 (Hz) at the vowel onset for four consonants, by vowel context and gender, averaged over positions and speakers.

Consonant	Gender	_a		_e		_i		_o		_u	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
s	f	2943	124	2892	181	2890	130	2947	134	3012	161
s <sup>j</sup>	f	2972	127	3063	133	3112	156	2707	155	2763	120
ʂ	f	2749	243	2826	155	2918	168	2728	220	2790	182
ʃ <sup>j</sup>	f	2844	228	3039	241	3102	211	2679	176	2691	148
s	m	2702	208	2628	117	2543	155	2746	214	2849	198
s <sup>j</sup>	m	2608	79	2600	127	2736	119	2548	210	2542	110
ʂ	m	2516	178	2461	133	2632	140	2545	250	2549	183
ʃ <sup>j</sup>	m	2606	141	2611	174	2754	87	2531	243	2517	211

**Figure A1** Plots showing mean COG at the midpoint (Hz) by mean F2–F1 difference at the vowel onset (Hz) for the fricatives /s/, /s<sup>j</sup>/, /ʂ/, and /ʃ<sup>j</sup>/ separately in each vowel context (averaged over positions and speakers).

## References

- Avanesov, Ruben I. 1948. *O dolgix shipjashchix v russkom jazyke* [About long hushing consonants in the Russian language] (Doklady i soobshchenija filologicheskogo fakul'teta MGU [Presentations and reports of the Philological Faculty of Moscow State University] 6).
- Avanesov, Ruben I. 1984. *Russkoe literaturnoe proiznoshenie* [Russian literary pronunciation]. Moscow: Prosveshchenie.
- Baudouin de Courtenay, Ivan A. 1964. Ob otnoshenii russkogo pis'ma k russkomu jazyku [About the relation between the Russian orthography and the Russian language]. *Izbrannye Trudy po obshchemu jazykoznaniju* [Selected works on Russian linguistics] 11. Moscow.
- Beckman, Mary E. & Jan Edwards. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In John Kingston & Mary E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the grammar and the physics of speech*, 152–178, Cambridge: Cambridge University Press.
- Bhat, D. N. S. 1978. A general study of palatalization. In Joseph H. Greenberg (ed.), *Universals of language*, vol. 2: *Phonology*, 47–91. Stanford, CA: Stanford University Press.



- Boersma, Paul & David Weenink. 2015. Praat: Doing phonetics by computer [Computer program]. <http://www.praat.org/>.
- Bolla, Kálmán. 1981. *A conspectus of Russian speech sounds*. Cologne: Böhlau.
- Bondarko, Liya V. & Ludmila A. Verbitskaya (eds.). 1987. *Interferentsija jazykovyx system* [Interference of sound systems]. St. Petersburg: Leningrad State University Publications.
- Borkovskii, Viktor I. & Petr S. Kuznetsov. 1965. *Istoricheskaia grammatika russkogo iazyka* [Historical grammar of the Russian language]. Moscow: Nauka.
- Carlton, Terence R. 1990. *Introduction to the phonological history of the Slavic languages*. Columbus, OH: Slavica.
- Clements, G. N. 2003. Feature economy in sound systems. *Phonology* 20, 287–334.
- Derkach Miron, Gunnar Fant & Antonio de Serpa-Leitao. 1970. *Phoneme coarticulation in Russian hard and soft VCV-utterances with voiceless fricatives* (Speech Research Quarterly Progress and Status Report (STL-QPSR) 2–3).
- Dmitrieva, Olga. 2017. Production of geminate consonants in Russian: Implications for typology. In Haruo Kubozono (ed.), *Aspects of geminate consonants*. Oxford: Oxford University Press.
- Fant, Gunnar. 1960. *Acoustic theory of speech production*. Mouton: The Hague.
- Flemming, Edward S. 2002. *Auditory representations in phonology*. New York: Routledge. [A published version of the 1995 Ph.D. dissertation, UCLA].
- Funatsu, Seiya & Shigeru Kiritani. 1998. Perceptual properties of Russians with Japanese fricatives. *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP)*, Sydney, Australia, 4 pp.
- Gordon, Matthew, Paul Barthmaier & Kathy Sands. 2002. A cross-linguistic acoustic study of fricatives. *Journal of the International Phonetic Association* 32, 141–174.
- Hamann, Silke. 2004. Retroflex fricatives in Slavic languages. *Journal of the International Phonetic Association* 34, 53–67.
- Hughes, George W. & Morris Halle. 1956. Spectral properties of fricative consonants. *The Journal of the Acoustical Society of America* 28(2), 303–310.
- Iskarous, Khalil & Darya Kavitskaya. 2010. The interaction between contrast, prosody, and coarticulation in structuring phonetic variability. *Journal of Phonetics* 38, 625–639.
- Jesus, Luis M. T. & Christine H. Shadle. 2002. A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics* 30(3), 437–464.
- Jones, Daniel & Dennis Ward. 1969. *The phonetics of Russian*. Cambridge: Cambridge University Press.
- Jongman, Allard, Rاتree Wayland & Serena Wong. 2000. Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America* 108, 1252–1263.
- Keating, Patricia. 1991. Coronal places of articulation. In C. Paradis & J.-F. Prunet (eds.), *The special status of coronals: Internal and external evidence* (Phonetics and Phonology 2), 29–48. San Diego, CA: Academic Press.
- Kedrova, Galina, Nikolai V. Anisimov, Leonid M. Zaharov & Yuriy A. Pirogov. 2008. Magnetic resonance investigation of palatalized stop consonants and spirants in Russian. *The Journal of the Acoustical Society of America* 123(5), 3325–3325.
- Kochetov, Alexei. 2006. Testing licensing by cue: A case of Russian palatalized coronals. *Phonetica* 63, 113–148.
- Kochetov, Alexei & Alevtina Lobanova. 2007. Komi-Permyak coronal obstruents: Acoustic contrasts and positional variation. *Journal of International Phonetic Association* 37, 51–82.
- Kochetov, Alexei & Milica Radišić. 2009. Latent consonant harmony in Russian: Experimental evidence for Agreement by Correspondence. In Maria Babyonyshev, Darya Kavitskaya & Jodi Reich (eds.), *Proceedings of the Seventeenth Formal Approaches to Slavic Linguistics (FASL) Meeting*, 111–130. Ann Arbor, MI: Jindřich Toman's Michigan Slavic Publications.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Wiley-Blackwell.
- Lee, Wai-Sum & Eric Zee. 2003. Illustrations of the IPA: Standard Chinese (Beijing). *Journal of the International Phonetic Association* 33, 109–112.
- Liljencrants, Johan & Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48(4), 839–862.

- Lindblom, Björn & Ian Maddieson. 1988. Phonetic universals in consonant systems. In Larry M. Hyman & Charles N. Li (eds.), *Language, speech and mind*, 62–78. London: Routledge.
- Litvin, Natallia. 2014. An ultrasound investigation of secondary velarization in Russian. M.A. thesis, University of Victoria.
- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
- Matusevich, Margarita I. 1976. *Sovremennyi russkii iazyk: fonetika* [The modern Russian language: Phonetics]. Moscow: Prosveshchenie.
- Munson, Benjamin, Elizabeth C. McDonald, Nancy L. DeBoe & Aubrey R. White. 2006. The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics* 34(4), 202–240.
- Nowak, Pavel M. 2006. The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of Phonetics* 34(2), 139–152.
- Ohala, John J. 1979. Phonetic universals in phonological systems and their explanation [Summary of symposium moderator's introduction]. *Proceedings of the 9th International Congress of Phonetic Sciences* (ICPhS IX), vol. 2, 5–8. Copenhagen: Institute of Phonetics.
- Padgett, Jaye & Marzena Żygis. 2007. The evolution of sibilants in Polish and Russian. *Journal of Slavic Linguistics* 15(2), 291–327.
- Purcell, Edward T. 1979. Formant frequency patterns in Russian VCV utterances. *The Journal of Acoustic Society of America* 66(6), 1691–1702.
- Scherba, Lev V. 1957. Teorija russkogo pis'ma [The theory of Russian orthography]. In Lev V. Scherba (ed.), *Izbrannye raboty po russkomu jazyku* [Selected works on Russian], 144–175. Moscow.
- Shupljakov, Victor, Gunnar Fant & Antonio de Serpa-Leitao. 1968. *Acoustical features of hard and soft Russian consonants in connected speech: A spectrographic study* (Speech Research Quarterly Progress and Status Report (STL-QPSR) 4).
- Skalozub, Larisa G. 1963. *Palatogrammy i rentgenogrammy soglasnykh fonem russkogo literaturnogo iazyka* [Palatograms and x-ray images of Russian consonants]. Kiev: Izdatel'stvo Kievskogo universiteta.
- Spinu, Laura, Irene Vogel & H. Timothy Bunnell. 2012. Palatalization in Romanian: Acoustic properties and perception. *Journal of Phonetics* 40, 54–66.
- Steriade, Donca. 1997. Phonetics in phonology: The case of laryngeal neutralization. Ms., UCLA.
- Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.
- Zinder, Lev R. 1963. *Fonematičeskaja sušh'nost' dolgogo palatalizavannogo [j':] v russkom jazyke* [Phonemic nature of the long palatalized [j':] in the Russian language] (*Filologičeskie nauki* [Philological sciences] 2).
- Zsiga, Elizabeth C. 2000. Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28(1), 69–102.
- Żygis, Marzena & Jaye Padgett. 2010. A perceptual study of Polish sibilants, and its implications for historical sound change. *Journal of Phonetics* 38(2), 207–226.