音声研究　第21巻第2号
2017(平成29)年8月
1–17頁

Journal of the Phonetic Society
of Japan, Vol. 21 No. 2
August 2017, pp. 1–17

研究論文

# Acoustic Manifestation of Russian Word-final Devoicing in Utterance-medial Position

Mayuki MATSUI*,**, Yosuke IGARASHI*** and Shigeto KAWAHARA****

非発話末におけるロシア語word-final devoicingの音響的実現

**SUMMARY:** Word-final devoicing is a common phonological process in many languages, including Russian. The current study examines word-final devoicing in utterance-medial position to tease apart word-final devoicing from utterance-final devoicing. Our new findings include: (i) neutralization is incomplete in utterance-medial position; and (ii) stops and fricatives differ in terms of implementation of devoicing. By comparing the current results with the previous studies, we entertain the possibility that phonetic implementations of incomplete neutralization may be sensitive to prosodic positions.

**Key words:** word-final devoicing, incomplete neutralization, manner of articulation, prosodic position, acoustic phonetics, Russian

## 1. Introduction

### 1.1 General Introduction

Many languages make lexical contrasts using the larynx, such as voicing and aspiration. They are sometimes collectively referred to as laryngeal contrasts. However, exactly how the larynx is used to make laryngeal contrasts is still to be fully understood (Ladefoged and Maddieson 1996, p. 49); also, another important issue that remains open is how laryngeal-based contrasts manifest themselves acoustically in different positions. The present study examines how a word-final—apparently "neutralized"—voicing contrast is implemented in utterance-medial position through a case study of Russian. The current study is situated within the larger enterprise of examining the presence and the nature of so-called "incomplete neutralization". Russian has a two-way voicing contrast (voiced vs. voiceless) in stops and fricatives. Previous phonological accounts have assumed that Russian voiced obstruents are fully devoiced in word-final positions, resulting in complete neutralization of the voicing contrast (e.g., Halle 1959, Hamilton 1980); that is, underlyingly voiced obstruents become identical to underlyingly voiceless obstruents by way of a categorical phonological process, and hence their phonetic realizations should be identical. However, later careful instrumental studies have shown that the Russian devoicing may be phonetically incomplete: small but consistent acoustic differences, which are demonstrably perceptible (e.g., Kharlamov 2012, 2015), have been observed between "neutralized" obstruents in Russian (among others, Dmitrieva et al. 2010, Kharlamov 2012, 2014, Kulikov 2012). This finding—that devoicing may not be complete—is actually reported in many other languages as well (Port and O'Dell 1985, Roettger et al. 2011, 2014 for German; Dinnsen and Charles-Luce 1984 for Catalan; Warner et al. 2004 for Dutch).

Indeed, an increasing number of acoustic and perceptual studies have documented the incomplete neutralization of word-final devoicing in Russian and other languages. However, most of these studies examined word-final devoicing in utterance-final position, thereby conflating "word-final" devoicing and "utterance-final" devoicing. In other words, it is still not clear whether incomplete neutralization occurs in word-final position in general, or incomplete neutralization occurs only in word-final and utterance-final position; i.e., word-final devoicing in utterance-medial position is understudied. As will be discussed in the following sections, utterance-final position may not be the ideal position to explore the nature of (incomplete)

* National Institute for Japanese Language and Linguistics（国立国語研究所）
** Japan Society for the Promotion of Science（日本学術振興会）
*** Hitotsubashi University（一橋大学）
**** The Institute of Cultural and Linguistic Studies, Keio University（慶應義塾大学言語文化研究所）

phonological word-final devoicing.

The present paper has two aims. One is to fill the above-mentioned descriptive gap by examining word-final devoicing in utterance-medial positions, which will advance our understanding of the nature of word-final devoicing and incomplete neutralization in general. Another is to examine word-final devoicing both in stops and fricatives, which enables us to address the question of whether stops and fricatives differ in terms of (in)complete neutralization. In addition, the current experiment attempts to minimize to the extent possible the effects of lexical factors and orthographic influence.

In the following sections, we will first review the previous studies of word-final devoicing and incomplete neutralization in general (section 1.2) and of Russian word-final devoicing in particular (section 1.3). The specific questions to be addressed in the present study are stated in section 1.4. Section 2 describes the method. Section 3 presents an acoustic analysis of word-final devoicing in utterance-medial position. Section 4 discusses the results of the acoustic analysis. Finally, section 5 concludes the paper.

### 1.2 Word-final Devoicing and Incomplete Neutralization

Word-final devoicing (henceforth, final devoicing) is a common phonological process, both synchronically and diachronically (Blevins 2006, Myers 2012). In many languages, such as Catalan, Dutch, German, and Russian, voiced obstruents devoice in domain-final position (see Myers 2012 for an extensive cross-linguistic overview). In languages with final devoicing, it has been traditionally formalized, especially within the generative phonology tradition, that the distinction between voiced and voiceless obstruents is completely lost in some domain-final position by way of a phonological rule. In this view, minimal-pairs such as /luk/ and /lug/ should result in complete homophones, [luk] (e.g., Akishina and Baranovskaja 2010, Cubberley 2002, Halle 1959, Hamilton 1980, Wade 2010)[1].

However, a number of experimental studies have found that neutralization may not be complete, contrary to the assumption held by these formal theories; that is, for example, [luk] derived from /luk/ is phonetically slightly different from [luk] derived from /lug/. Throughout this paper, we use the term “incomplete (neutralization)” if there is an observable acoustic difference between segments that are claimed to be neutralized.

While there is now a large number of studies which demonstrate that some neutralization patterns can be incomplete, some researchers have argued that incomplete neutralization is nothing more than an experimental artifact, such as orthographic influence (e.g., Fourakis and Iverson 1984). Therefore, it is important to control for as many confounding factors as possible, in order to establish the existence of incomplete neutralization in Russian and other languages. This task is especially important because incomplete neutralization is not predicted by the standard theories of phonology: if a contrast is neutralized in phonology, the outputs of phonology, or the inputs to phonetics, should be identical. Therefore, phonetics should not treat the neutralized pair of sounds differently. See Port and Leary (2005) for extensive discussion on the problems that incomplete neutralization presents to formal phonological theories[2].

A classic study of incomplete neutralization is on German final devoicing (Port and O’Dell 1985). Port and O’Dell (1985) examined German final devoicing using 10 minimal pairs of words produced in isolation for stop consonants. Their results showed that there were significant durational differences between the underlyingly voiced stops and the underlyingly voiceless stops. For example, the vowels preceding an underlyingly voiced stop (e.g., *Rad* [ʁaːt] ‘wheel’) were on average 15 ms longer than those preceding an underlyingly voiceless stop (e.g., *Rat* [ʁaːt] ‘council’). Further, those differences were perceptible to native German listeners, although the accuracy rate was only slightly above chance-level. Based on these findings, Port and O’Dell (1985) concluded that neutralization was incomplete. Similar patterns of incomplete neutralization have since been identified in many languages, in which small but consistent phonetic differences have been observed between segments—or supra-segmental contrasts—that are assumed to be completely neutralized (among others, Charles-Luce 1985, Piroth and Janker 2004, Port and O’Dell 1985, Port and Crawford 1989, Roettger et al. 2011, 2014 for German final devoicing; Dinnsen and Charles-Luce 1984 for Catalan final devoicing; Ernestus and Baayen 2006, 2007, Warner et al. 2004 for Dutch final devoicing; Dmitrieva et al. 2010, Kharlamov 2012, 2014, Kulikov 2012 for Russian final devoicing; Braver 2013, Herd et al. 2010, for American English flapping; Fourakis and Port 1986 for American English stop epenthesis; Yu 2007 for Cantonese tonal merger).

There has now thus been a large body of work claiming the existence of incomplete neutralization. However, a number of phonetic experiments have also found that our speech production patterns can be influenced by multiple factors, which may raise the question of whether the previous findings on incomplete neutralization could be an experimental artifact arising from these factors. For example, the extent to which some contrasts are incompletely neutralized has been shown to be task-dependent, including the presence/absence of orthographic stimuli (Fourakis and Iverson 1984, Kharlamov 2012, 2014, Warner et al. 2006). Fourakis and Iverson (1984) examined the influence of orthographic stimuli on the acoustic realization of final obstruent devoicing in German. Their results showed that, when the speakers were presented with orthographic stimuli (i.e., a traditional reading task), there was clear evidence for incomplete neutralization. In contrast, no clear evidence for incomplete neutralization was found when the speakers were asked to conjugate verbs without orthographic stimuli (a verbal conjugation task). On the basis of their results, Fourakis and Iverson (1984) concluded that the apparent case of incomplete neutralization found by Port and O'Dell (1985) was merely due to the effect of the German orthography.

Despite the concern raised by Fourakis and Iverson (1984), recent studies suggest that the phonetic differences in a neutralizing position cannot be attributed solely to orthographic influence. In their study on German final devoicing, Roettger et al. (2011, see also 2014) adopted an experimental task in which nonce words were elicited by presenting the stimuli to the speakers with auditory cues, thereby minimizing the effects of orthography. The results showed that significant differences were observed in terms of preceding vowel duration and burst intensity. They concluded that, contrary to Fourakis and Iverson (1984), incomplete neutralization does occur even when the speakers are not exposed to orthographic differences. The orthographic effects on final devoicing have been examined in other studies on incomplete neutralization as well (Ernestus and Baayen 2006, Kharlamov 2012, Warner et al. 2006), and these studies showed that, in general, acoustic differences are greater when the speakers are presented with orthographic cues than when they are not, but that nevertheless, incomplete neutralization is observed even without the presence of orthographic input.

In addition to the effect of orthography, word-specific factors such as lexical frequency (Pierrehumbert 2002, VanDam and Port 2005), and the presence/absence of a lexical neighbor (Baese-Berk and Goldrick 2009, Wright 2004) are shown to affect phonetic realizations, potentially resulting in apparent cases of incomplete neutralization. For example, VanDam and Port (2005) showed that the VOT of English /t/ is longer in low lexical frequency words than in high frequency words. Baese-Berk and Goldrick (2009) found that the presence/absence of minimal pairs among lexical items can affect the duration of VOT in English. They showed that the VOT of a word-initial /p/ was significantly longer when the word in question has a minimally contrasting neighbor (e.g., *pox* [pʰɔks] which has an existing minimal pair, *box* [bɔks]) than when it does not have such a neighbor (e.g., *posh* [pʰɔʃ] vs. **bosh* [bɔʃ]).

These lexical effects on phonetic implementation should be taken into consideration in the study of incomplete neutralization. For example, suppose that in studying a case of incomplete neutralization, we may find that the burst of [t] of the underlying /t/ is longer than that of underlying /d/; however, that longer burst may be due to the fact that the words containing /t/ happen to be less frequent or have a minimal pair (i.e., hyperarticulated due to lexical reasons). Similarly, suppose that words containing /d/ were hyperaticulated due to some lexical factor. This would result in longer preceding vowel duration, which may look like incomplete neutralization. Thus, it is necessary to show that incomplete neutralization occurs independently of such lexical frequency effects. At the very least, in studying incomplete neutralization, it would be ideal to minimize the effects of lexical factors and orthographic influences, to the extent possible.

### 1.3 Background on Russian Word-final Devoicing

Russian is one of the many languages in which the distinction between underlyingly voiced and voiceless obstruents is thought to be neutralized in word-final position (e.g., Akishina and Baranovskaja 2010, Cubberley 2002, Hamilton 1980, Halle 1959, Wade 2010). The domain of final devoicing in Russian is considered to be a prosodic word, which includes prepositions and lexical content words (Padgett 2012). At the right edge of a prosodic word, underlyingly voiced and voiceless obstruents are, impressionistically speaking, homophonous (e.g., [luk] from /luk/ 'onion$_{\text{.MASC.NOM.SG.}}$' vs. [luk] from /lug/ 'meadow$_{\text{.MASC.NOM.SG.}}$'). However, the actual status of this neutralization has long been debated.

Previous instrumental studies have observed small acoustic differences between these neutralized "homophones" (Chen 1970, Dmitrieva et al. 2010, Kharlamov 2012, 2014, Kulikov 2012, Pye 1986, Shrager 2012). Using a word-reading task, Dmitrieva et al. (2010) showed that there are significant acoustic differences between underlyingly voiced obstruents and underlyingly voiceless obstruents in word-final position, when the speakers read the words in isolation. For monolingual speakers of Russian, constriction duration was significantly longer for underlyingly voiceless obstruents than for underlyingly voiced obstruents. They also found that the duration of the release burst of the underlyingly voiceless stop was significantly longer than that of the underlyingly voiced stops. Shrager (2012) examined dental stops in utterance-final position, and reported results that are similar to those of Dmitrieva et al. (2010). Chen (1970) and Kulikov (2012) reported that vowels preceding underlyingly voiced stops were longer than those preceding underlyingly voiceless stops. Kulikov (2012) also found that the first formant (F1) was lower before underlyingly voiced stops than before underlyingly voiceless stops. The results of these studies suggest that neutralization may be incomplete in Russian word-final position. However, the existence of incomplete neutralization in Russian can and should be more firmly established, because several factors that might affect phonetic details, discussed in section 1.2, were not thoroughly controlled for in those studies. Since most of the previous experiments involve Russian orthography, the "incomplete neutralization" found in those studies may have arisen due to the effect of orthography. Other studies have not controlled factors such as the lexical frequencies of the stimuli, which can lead to an apparent case of incomplete neutralization.

In his extensive study of methodological influence on final devoicing in Russian, Kharlamov (2012, 2014) found significant differences in closure duration and burst duration of stops, and frication duration of fricatives between underlyingly voiceless and voiced obstruents. Kharlamov (2012, 2014) examined Russian final devoicing with two tasks: a reading task and a picture-naming/fill-in-blank task, in which speakers were asked to produce existing Russian words in isolation. Since the latter task involved no orthographic stimuli, the observed differences cannot readily be attributed to orthographic effects. Moreover, regardless of whether or not the speakers were exposed to both members of lexical minimal pairs in the stimuli, neutralization was incomplete. In summary, Kharlamov (2012, 2014), arguably the best controlled experiment on Russian incomplete neutralization, found that the observed acoustic differences cannot be relegated away as methodological artifacts. Further, Kharlamov (2012) controlled for the lexical token frequencies of the target words, minimizing the potential effects of lexical factors discussed above. Yet, Kharlamov's (2012, 2014) studies targeted only utterance-final positions, which could be a concern, as further discussed in section 1.4.

To summarize, previous instrumental studies suggest that Russian final devoicing is demonstrably incomplete, and that acoustic differences found between supposedly neutralized obstruents cannot be regarded as the consequence of orthography, the presence or absence of a minimal pair in the lexicon, or lexical token frequency[3)]. We nevertheless feel that there are still some gaps to fill.

### 1.4 The Questions and the Motivations of This Study

While previous studies extensively documented Russian final devoicing in utterance-final position (e.g., Khalramov, 2012 2014), word-final devoicing in *utterance-medial* position is understudied. The only study that we are aware of is Kulikov (2012), who examined final devoicing in utterance-medial position by embedding the target words in a frame sentence. As discussed above, Kulikov's results showed that, for stops, the underlying voicing was distinguished by the preceding vowel duration and F1 before the stop, when the speakers performed a reading task. However, this study has a number of limitations. First, as discussed above, Kulikov (2012) used a reading task, and hence the results may have been due to orthographic influences. Second, since the study examined only a small set of minimal pairs of existing words, the results can be attributed to hyperarticulation caused by lexical factors. Third, since the focus of Kulikov's (2012) study was stop consonants, the behavior of fricatives in utterance-medial position is understudied.

There are some reasons to expect that final devoicing can be different in utterance-final position and in utterance-medial position. First, devoicing is phonetically motivated in utterance-final position (e.g., Myers 2012, Myers and Padgett 2014, Smith 1997, 2003), because utterance-final segments often suffer from passive devoicing. In non-speech mode, the vocal folds are separated apart to facilitate breathing; therefore, toward the end of an utterance, speakers tend to start spreading apart their vocal folds, resulting in cessation of vocal fold vibration. In addition to this "coarticulation to the non-speech state", the subglottal air pressure that is necessary to produce vocal fold vibration decreases over the course of an utterance, which also makes the utterance-final segments likely to devoice (see Myers and Padgett 2014, p. 400, for a recent summary and list of references on these issues; see also Barnes 2006 for phonetic and phonological characteristics of utterance-final position in general). Therefore, phonetic, physiologically-motivated utterance-final devoicing is commonly observed in languages even without phonological devoicing. On the other hand, there is no apparent phonetic motivation for devoicing in utterance-medial position if the word-final obstruent is followed by a vowel or a sonorant (again, see Myers and Padgett 2014). In this sense, phonological devoicing in utterance-final position may be confounded by the phonetic utterance-final devoicing; it is therefore important to investigate the behavior of word-final devoicing in non-utterance-final positions, especially when the target consonants are followed by vowels or sonorants.

In addition, since phonetic implementation patterns are significantly affected by differences in prosodic position, studying the difference between word-final position and utterance-final position in terms of final devoicing is in and of itself important. Deciphering the domain of phonological processes has been an issue with continued interests throughout the history of phonological theory (Selkirk 1980 *et seq*.), and how prosodic structures manifest themselves in terms of acoustics and articulation has also been a topic for intensive phonetic investigation (e.g., Byrd et al. 2006, Kuzla and Ernestus 2011, Wightman et al. 1992). The current study can be situated as a case study of this tradition as well.

Another reason to study incomplete neutralization in utterance-medial position is as follows: utterance-final segments are often lengthened, which is known as "final lengthening" (e.g., Klatt 1976, Wightman et al. 1992). This raises a potential concern for the study of incomplete neutralization: it may be the case that incomplete neutralization occurs only when the segments are lengthened. For example, speakers may implement the lengthening of the vowel preceding /d/, only when they have time to implement that gesture thanks to extra time provided by utterance-final lengthening. Alternatively, it could be the case that the durational difference increases proportionally, as segments are lengthened by utterance-final lengthening, resulting in apparent incomplete neutralization. Therefore, a question arises as to whether incomplete neutralization is observed, even in the context where utterance-final lengthening is irrelevant.

In summary, studying the phonetic implementation pattern of devoiced obstruents in word-final, utterance-medical position is important for several reasons.

Additionally, this study addresses another topic that has not been sufficiently addressed in the previous research: the difference between stops and fricatives. Stops and fricatives might differ in how they devoice, since voicing in stops and fricatives face different kinds of aerodynamic challenges (e.g., Davidson 2016, Ohala 1983, Solé 2002, Żygis et al. 2012). In stops, maintaining voicing during closure is challenging, because the rise in intraoral air pressure, which results from stop closure, inhibits glottal airflow. Speakers often deal with this problem by expanding their oral cavity, thereby lowering the intraoral air pressure. However, this simple solution does not work for voiced fricatives, because fricatives require high intraoral air pressure to cause frication in the first place (Ohala 1983, Solé 2002). Therefore, voicing in fricatives requires more delicate articulatory maneuvers compared with voicing in stops. Despite this difference, previous studies on incomplete neutralization usually focused on stops (e.g., Kulikov 2012 for Russian, Roettger et al. 2011, 2014 for German; Warner et al. 2004 for Dutch), and a direct comparison between stops and fricatives is rarely done. As mentioned in section 1.3, Dmitrieva et al. (2010) and Kharlamov (2012, 2014) tested both stops and fricatives in utterance-final position. Again, since utterance-final position is not an ideal condition to look at phonological final devoicing, it still remains unclear whether stops differ from fricatives in terms of final devoicing. Thus, it would be of interest to investigate how stops and fricatives may (or may not) differ in terms of how they devoice incompletely.

To summarize, the present study examines incomplete word-final devoicing from a few new perspectives through a case study of Russian word-final devoicing in utterance-medial position for both stops and fricatives. The specific questions to be addressed in this study are: (i) will we observe incomplete neutralization in utterance-medial position? and (ii) do stops and fricatives behave the same way in terms of phonological devoicing? To this end, we adopt the experimental paradigm used in Roettger et al. (2011, 2014) for German final devoicing to examine Russian word-final devoicing in utterance-medial position.

Throughout this paper, phonemic representations follow the latest IPA illustrations of Russian (Yanushevskaya and Bunčić 2015). For the sake of typographical simplicity, we use [t] and [d] without the diacritic to represent dental stops.

## 2. Methods

### 2.1 Participants

Twelve native speakers of Russian, aged 19–22 (Mean age: 20.8, SD: 0.9), participated in the production experiment. Seven speakers were female. All of the participants were born and raised around the city of Orenburg in Russia. The variety of Russian spoken in this region is reportedly not different from Standard Russian (“*literaturnyj Russkij jazyk*”) in terms of voicing contrast and final devoicing (Avanesov and Orlova 1965, p. 84). None of the participants reported that they use other languages on a daily basis. Two of the participants reported that their national background is both Russian and Kazakh, although they reiterated that they could only speak Russian. All participants were thus monolingual speakers of Russian.

## 2.2 Speech Materials

All the non-palatalized obstruents that have a voicing contrast in Russian were examined: bilabial (/p/, /b/), dental (/t/, /d/), and velar (/k/, /g/) stops and labiodental (/f/, /v/), alveolar (/s/, /z/) and post-alveolar (/ʃ/, /ʒ/) fricatives. The vowels preceding the target obstruents were /i/, /e/, /a/, /o/ or /u/. 30 minimal pairs (6 consonant types×5 vowels) were used as the target items. The target items were all nonce nouns. The current study used nonce words in order to minimize the word-specific phonetic effects, such as lexical frequency effects or the effect of the existence of a minimal pair as reviewed in section 1. To create nonce words, we consulted a native speaker of Russian. For some speakers, a certain number of words could potentially be interpreted as existing words, because different speakers have a different lexicon. At the very least, the post-experimental interview confirmed that the majority of our participants interpreted the speech material as non-existing words. We also calculated the nonce words’ lexical neighborhood densities, since their lexical neighbors could have been activated in the elicitation phase[4)]. We therefore provide a brief report of lexical neighborhood structure around the nonce words in the Appendix, which used substitution neighborhood density calculation (SND; i.e., the number of lexical items which differ in one segment only). The SNDs were calculated using an online dictionary based on the Russian National Corpus (Lyashevskaya and Sharov 2009) and Phonological CorpusTools 2.1.0 (Hall et al. 2016).

All of the target items were monosyllabic words, consisting of either CVC or CCVC structures. The list of all the target words appears in the Appendix. In addition to the target items, 60 fillers were added.

In order to elicit these target items without orthography, we used auditory prompts. The auditory prompts were produced by a native speaker of Russian (a female speaker in her twenties from Orenburg). As described in the next section, the stem-final obstruents used in these auditory prompts were followed by a vowel; hence the voicing contrast was not neutralized. The acoustic properties of the auditory prompts are provided in Table 1. The overall patterns between voiceless and voiced obstruents are consistent with those that are reported cross-linguistically (Kingston and Diehl 1994, among others): Preceding vowels are longer before voiced obstruents, closure voicing is longer during voiced obstruents, constriction durations are shorter during voiced obstruents, release bursts are longer for voiceless obstruents, and F0 and F1 are generally higher preceding voiceless obstruents.

## 2.3 Procedures

The elicitation method followed that of Roettger et al. (2011, 2014). Within each trial, one auditory prompt was first presented to the participants via headphones (SONY MDR-Z700). The auditory prompts were embedded in a sentence such as: /v kafe sʲidʲelʲi pʲatʲ X-ov/ (‘In the café five $X_{\text{-MASC.GEN.PL.}}$ were sitting’, where X is a target). The auditory prompts were followed by a masculine genitive plural ending /ov/ [əf], and hence the target words in the prompts were placed in a *non*-neutralizing position. A sample auditory prompt is given in (1). A perception test confirmed that the voicing contrast in the auditory prompt was very easily perceived by the native listeners of Russian (10 listeners, forced-choice identification task, average *d′*: 4.135, SD: 0.2)[5)].

Next, in response to each auditory prompt, the participants were asked to produce the target word with the fixed frame sentence such as: /odʲin X jel pʲetʃʲenʲje/ (‘One $X_{\text{.MASC.NOM.SG.}}$ was eating a cookie.’). They were presented with this frame sentence before the experimental session. In the frame sentence, the participants were expected to use a nominative singular form, where the target obstruent, because it is word-final, is located in a neutralizing position. The word-final obstruent was always followed by a sonorant-onset verb /jel/ [jel] ‘was eating’[6)]. A sample sentence is shown in (2).

(1) Auditory prompt:
/v kafe sʲidʲelʲi pʲatʲ <u>prʲigov</u>/
‘In the café five $\text{prig}_{\text{-MASC.GEN.PL.}}$ were sitting.’ (The target nonce noun: ‘prig’ /prʲig/)

(2) Expected response:
/odʲin <u>prʲig</u> jel pʲetʃenʲje/
‘One $\text{prig}_{\text{.MASC.NOM.SG.}}$ was eating a cookie.’

**Table 1 Summary of the acoustic properties of the auditory prompts (a female speaker, non-neutralizing position, 30 stops, 30 fricatives). Average acoustic values with one standard deviation (in parentheses). The first four parameters are all duration. See section 2.4 for details.**

| Stops | Bilabial | | Dental | | Velar | |
|---|---|---|---|---|---|---|
| | Voiceless /p/ | Voiced /b/ | Voiceless /t/ | Voiced /d/ | Voiceless /k/ | Voiced /g/ |
| *Vowel* [ms] | 114 (7) | 123 (17) | 110 (30) | 124 (18) | 117 (15) | 131 (16) |
| *Voicing* [ms] | 16 (4) | 71 (14) | 12 (6) | 61 (5) | 12 (3) | 58 (5) |
| *Constriction* [ms] | 92 (15) | 71 (14) | 97 (5) | 61 (5) | 80 (10) | 58 (5) |
| *Release* [ms] | 14 (3) | 6 (2) | 16 (3) | 7 (2) | 25 (4) | 12 (4) |
| *F0 pre C* [Hz] | 196 (5) | 197 (12) | 195 (6) | 193 (8) | 190 (4) | 187 (3) |
| *F1 pre C* [Hz] | 461 (94) | 474 (88) | 473 (177) | 491 (124) | 462 (105) | 415 (67) |

| Fricatives | Labiodental | | Alveolar | | Post alveolar | |
|---|---|---|---|---|---|---|
| | Voiceless /f/ | Voiced /v/ | Voiceless /s/ | Voiced /z/ | Voiceless /ʃ/ | Voiced /ʒ/ |
| *Vowel* [ms] | 120 (18) | 138 (13) | 121 (25) | 138 (20) | 123 (10) | 174 (27) |
| *Voicing* [ms] | 12 (2) | 53 (8) | 10 (4) | 73 (14) | 17 (6) | 64 (7) |
| *Constriction* [ms] | 120 (13) | 53 (8) | 137 (13) | 80 (12) | 142 (15) | 64 (7) |
| *F0 pre C* [Hz] | 194 (8) | 193 (9) | 193 (4) | 185 (6) | 197 (7) | 191 (6) |
| *F1 pre C* [Hz] | 528 (169) | 471 (92) | 427 (74) | 406 (55) | 497 (211) | 418 (47) |

The 60 target items and 60 fillers were presented in a pseudo-random order, separated in four blocks: minimal pairs never appeared within the same block or adjacent blocks. Orders of items within each block were randomized. Stimuli were presented with Praat (Boersma and Weenink 2010). Before the experiment, the participants performed three practice trials with existing words and seven trials with pseudo words. The participants were encouraged to imagine that they would continue a story according to the Russian sentence (i.e., the auditory prompt) that they heard.

The recordings were conducted in a quiet room in Orenburg State University and the university dormitory. The tokens were recorded using a portable recorder (SONY PCM-M10) with a stereo microphone (SONY ECM-MS907) with a 44.1 kHz sampling rate at a 16 bit quantization level. The first channel of the stereo recordings was extracted for the acoustic analysis.

### 2.4 Acoustic Analysis

The target words were acoustically analyzed using the Praat speech analysis software (Boersma and Weenink 2010). The following temporal and spectral properties, which are associated with a voicing contrast (e.g., Kingston and Diehl 1994), were measured:

1. Duration of preceding vowel (*Vowel*);

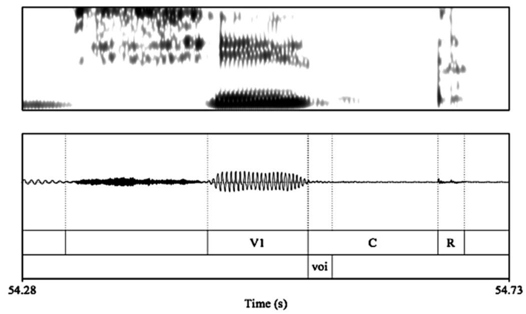


**Figure 1 A representative spectrogram and waveform (/zʲid/).**

2. Voicing duration during closure/frication (*Voicing*);
3. Constriction (closure or frication) duration (*Constriction*);
4. Release duration of stop consonants (*Release*);
5. F0 at the edge of the preceding vowel *(F0 pre C*);
6. F1 at the edge of the preceding vowel (*F1 pre C*).

The representative waveforms and spectrograms are given in Figure 1 to illustrate the measurement protocol. For *Vowel*, when a nasal, lateral or fricative preceded the vowel, the onset of the vowel was defined as the time point of abrupt change in amplitude of the waveform. When a trill preceded the target vowel, the vowel onset was defined as the end of a short silence. The offset of the vowel was identified based on an abrupt change in amplitude and the shape of the waveform.

Voicing duration during closure/frication was defined as the interval between the offset of the preceding vowel and the end of periodicity in the waveform. The presence of low frequency energy in the spectrogram was also consulted when the end of vocal fold vibration cannot be unambiguously determined based on the waveform. For *Constriction*, closure duration was measured from the vowel offset to the onset of the burst noise. Frication duration was measured from the vowel offset until the end of fricative noise with an abrupt change in the amplitude in the spectrogram. For *Release*, the release duration of stops was measured from the onset of the burst until the end of noise or an abrupt change in the waveform and spectrogram. For *F1 and F0 pre C*, F1 at the offset of the preceding vowel (i.e., the onset of *C*) was extracted by using the Burg algorithm of Praat. F0 was measured using the autocorrelation algorithm.

## 2.5 The Data Set

In total, 720 tokens were elicited from the speakers (60 tokens×12 speakers). 28 tokens (3.8%) were discarded because of pronunciation errors or recording errors. The remaining 692 tokens were submitted for the acoustic analyses.

In the course of acoustic analysis, some tokens were excluded due to the difficulty of unambiguously locating the segmental boundary. The tokens in which the F0 was not correctly extracted due to irregular vocal fold vibration were excluded from the F0 analysis. Two pairs (/s$^j$es/-/s$^j$ez/ and /dn$^j$ek/-/dn$^j$eg/) were systematically excluded from the F1 analyses since some participants produced the onset consonant with palatalization, while others without, which can potentially affect F1. The final number of tokens included for statistical analysis is summarized in Table 2.

**Table 2 Number of tokens submitted to statistical analysis.**

| *Measurement* | Stop | Fricative |
|---|---|---|
| *Vowel* | 340 | 352 |
| *Voicing* | 336 | 352 |
| *Constriction* | 336 | 352 |
| *Release* | 334 | N.A. |
| *F0 pre C* | 318 | 332 |
| *F1 pre C* | 318 | 330 |

## 2.6 Statistical Analysis

We fit linear mixed-effect regression models for each measurement by using lme4 package (version 1.1–7, Bates et al. 2014) in *R* 3.1.0 (R Core Team 2014). In our models, Underlying Voicing (Uvoi) and Place of Articulation (POA) were coded as fixed effects, and participants and items as random effects; both a random intercept and slope were included in the model (cf. Barr et al. 2013). Random slopes for items were not coded for POA, since POA are not random across items in the current design. While Uvoi was dummy-coded (as default setting), POA was sum-coded, so that the intercepts in our model were centered to “devoiced” level averaged across three places of articulation[7)].

In each model, *p*-values are estimated by likelihood-ratio tests. Since separate acoustic dimensions were assessed at the same time, the $\alpha$-level was adjusted for multiple comparisons by using Bonferroni correction: $\alpha$ adjusted=0.008 (0.05/6) for stops, $\alpha$ adjusted=0.01(0.05/5) for fricatives.

## 2.7 Predictions

As discussed in section 1.4, the specific questions addressed in this study are: (i) when we minimize orthographic and lexical factors, is neutralization still incomplete? (ii) will we observe incomplete neutralization in utterance-medial position? (iii) do stops and fricatives behave the same way in terms of devoicing?

If there is no such thing as incomplete neutralization, beyond experimental artifacts reviewed in section 1, then the prediction is that there should be no differences between underlyingly voiced obstruents and voiceless obstruents, in any of the acoustic dimensions studied here.

On the other hand, as reviewed in section 1.3, Kulikov (2012), who tested the same prosodic environment as the current study, showed that, for stops, the underlying voicing was distinguished by the preceding vowel duration and F1 before the obstruent when the speakers performed a reading task.

If the underlyingly voiceless and “devoiced” stops were in fact incompletely neutralized as reported in Kulikov (2012), we would expect a difference in the preceding vowel duration of stops and/or F1 before final stops in the current experiment. If stops and fricatives should behave identically in terms of devoicing, we would observe a difference in the vowel preceding a final fricative and/or F1 before final fricatives, just like final stops.

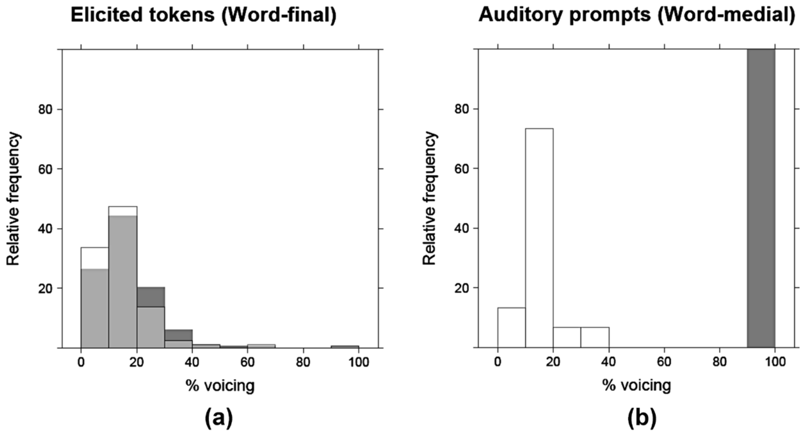


**Figure 2 The distribution of closure voicing (%) (10% bin) in word-final position (a) and in non-word-final position (b). Shaded bars represent underlyingly voiced tokens; white bars represent underlyingly voiceless tokens. The lighter gray represents the overlap between the two categories.**

## 3. Results

### 3.1 Stops

The word-final stops in the dataset were typically realized with short vocal fold vibration during stop closure. For the sake of explicit comparison, the distribution of closure voicing (%) is compared between word-final position (Figure 2(a)) (the experimental targets) and non-word-final position (Figure 2(b)) (the auditory prompts). As can be seen in Figure 2(b), the distributions of underlyingly voiceless and underlyingly voiced stops do not overlap in non-word-final position, which shows that the voicing contrast was maintained in the auditory prompts. On the other hand, the distributions of underlyingly voiceless and underlyingly voiced stops entirely overlap in word-final position. Closure voicing (%) of the underlyingly voiced stops was significantly different between word-final position and non-word-final position [independent two-group Mann-Whitney U Test, $W=7.5$, $p<0.001$], whereas that of the underlyingly voiceless stops was not [$W=1047$, $p=0.2942$]. We take this to be evidence that word-final stops were phonologically devoiced, despite being followed by a sonorant. The mean voicing proportion during stop closure was 16% (SD=12). The representative tokens of word-final stops are illustrated in Figure 3(a) and 3(b).

Having established that word-final devoicing does take place, Table 3 shows the average acoustic properties of the target obstruents, broken down by underlying voicing (voiceless and its "devoiced" counterparts) and place of articulation for each phonetic dimension.

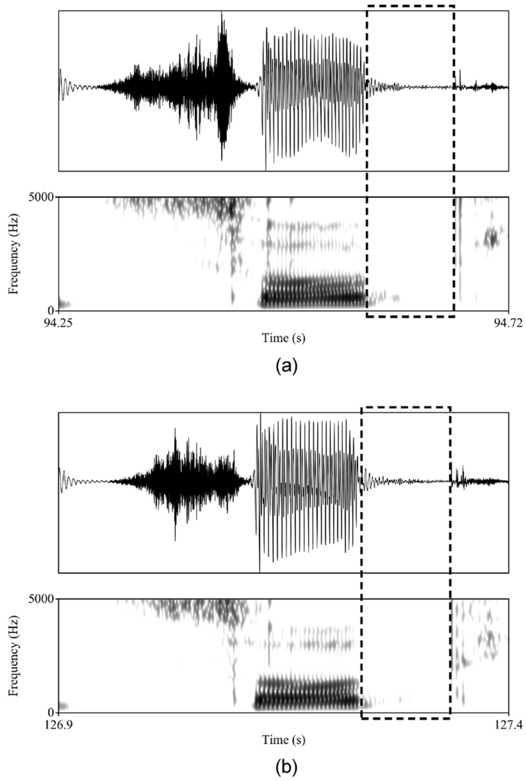


**Figure 3 (a) A representative token of word-final voiceless stops (/tsot/). The closure interval is indicated by a box with a broken line. (b) A representative token of word-final devoiced stops (/tsod/). The closure interval is indicated by a box with a broken line.**

**Table 3 Average acoustic values with standard deviations (in parentheses) for measurements (stops). The measurement with statistical significance is indicated in boldface.**

| | Bilabial | | Dental | | Velar | |
|---|---|---|---|---|---|---|
| | Voiceless /p/ | Devoiced /b/ | Voiceless /t/ | Devoiced /d/ | Voiceless /k/ | Devoiced /g/ |
| ***Vowel* [ms]** | **106 (18)** | **112 (17)** | **106 (25)** | **111 (21)** | **106 (19)** | **115 (23)** |
| *Voicing* [ms] | 13 (5) | 16 (7) | 15 (11) | 17 (13) | 14 (8) | 16 (9) |
| *Constriction* [ms] | 115 (20) | 114 (22) | 93 (20) | 95 (23) | 92 (20) | 92 (18) |
| *Release* [ms] | 54 (29) | 47 (25) | 52 (21) | 48 (24) | 61 (25) | 64 (31) |
| *F0 pre C* [Hz] | 239 (72) | 238 (76) | 238 (70) | 236 (69) | 241 (74) | 244 (75) |
| *F1 pre C* [Hz] | 453 (108) | 454 (104) | 450 (106) | 445 (108) | 428 (123) | 422 (104) |

To examine the effect of underlying voicing (Uvoi) on each measurement, we first compared two models: one with the interaction between Uvoi and POA and the other without. Likelihood ratio tests showed that there were no significant interactions between Uvoi and POA for any of the measurements [For *Vowel*, $\chi^2(2)=1.8024$, $p=0.4061$; For *Voicing*, $\chi^2(2)=0.4001$, $p=0.8187$; For *Constriction*, $\chi^2(2)=0.3746$, $p=0.8292$; For *Release*, $\chi^2(2)=3.4522$, $p=0.178$; For *F0 pre C*, $\chi^2(2)=1.8909$, $p=0.3885$; For *F1 pre C*, $\chi^2(2)=0.3143$, $p=0.8546$]. Therefore, we interpreted the results of the simpler model without the interaction term.

We found a significant effect of Uvoi in the duration of the preceding vowel [$\chi^2(1)=10.616$, $p<0.008$]. As can be seen in Table 3, the vowels preceding “devoiced” stops were longer than those preceding underlyingly voiceless stops. The regression coefficient ($\beta$) for Uvoi was -6.5 ms. Since the intercept (i.e., the reference level) in our model was centered to the “devoiced” level across POA, the model shows that the vowels preceding “devoiced” stops are 6.5 ms longer than those preceding voiceless stop. The effect of POA was not significant [$\chi^2(2)=0.1767$, $p=0.9154$]. A summary of the model for the preceding vowel duration is shown in Table 4.

For all other measurements (*Voicing*, *C*, *Release*, *F0 pre C and F1 pre C*), the effect of Uvoi did not reach significance. That is, we found evidence for incomplete neutralization, but only in terms of preceding vowel duration for stops.

### 3.2 Fricatives

Similar to stops, the word-final fricatives in the dataset were typically realized with short vocal fold vibration during frication. The distribution of voicing during fricative intervals is compared between word-final position (Figure 4(a)) and non-word-final position (Figure 4(b)).

**Table 4 A model summary for preceding vowel duration (for stops). Intercept is centered to “devoiced” level across POA. *p*-values smaller than 0.008 (after being adjusted for multiple comparisons, by Bonferroni correction) are indicated with “*”.**

| *Predictor* | Estimate ($\beta$) | Standard Error | t-value | |
|---|---|---|---|---|
| *Intercept (“devoiced”)* | 112.8652 | 4.0626 | 27.782 | |
| *Uvoi (voiceless)* | −6.4915 | 1.6867 | −3.849 | * |
| *POA (dental)* | −0.9177 | 4.427 | −0.207 | n.s. |
| *(velar)* | −1.0598 | 4.3669 | −0.243 | n.s. |

In the neutralizing environment, the two distributions overlap, suggesting that devoicing took place for fricatives word-finally. Voicing (%) of the underlyingly voiced fricatives was significantly different between word-final position and non-word-final position [independent two-group Mann–Whitney U Test, $W=1$, $p<0.001$], whereas that of the underlyingly voiceless fricatives was not [$W=1006$, $p=0.1271$]. We again take these to be evidence that word-final fricatives were phonologically devoiced, despite being followed by a sonorant. The mean voicing proportion during frication in word-final position was 10% (SD=6). Representative tokens are illustrated in Figures 5(a) and 5(b). Table 5 summarizes the acoustic properties of target words in the neutralizing environment.

We first compared two models, one with interaction between Uvoi and POA and one without. Again, the likelihood ratio tests showed no significant interactions for any of the measurements [For *Vowel*, $\chi^2(2)=0.6218$, $p=0.7328$; For *Voicing*, $\chi^2(2)=1.8266$, $p=0.4012$; For *Constriction*, $\chi^2(2)=1.0244$, $p=0.5992$; For *F0 pre C*, $\chi^2(2)=0$, $p=1$; For *F1 pre C*, $\chi^2(2)=0.4924$, $p=0.7818$]. Therefore, we interpret the results of the simpler model

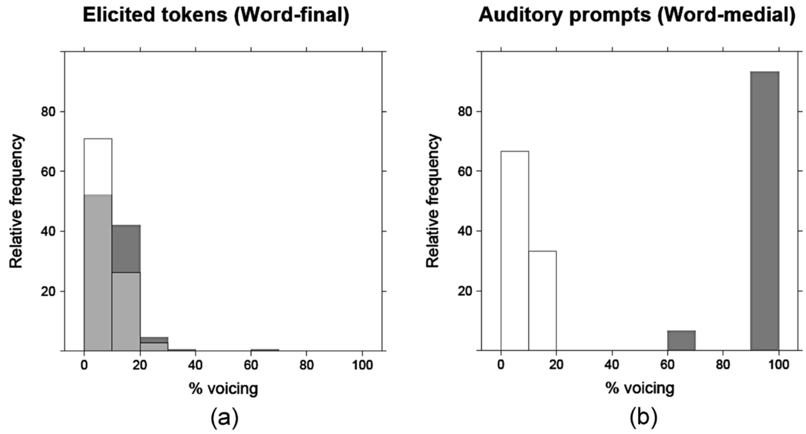


**Figure 4** **The distribution of voiceless and (de-)voiced fricatives along voicing proportion continuum (10% bin) in word-final position (a) and in non-word-final position (b). Shaded bars represent underlyingly voiced tokens; white bars represent underlyingly voiceless tokens. The lighter gray represents the overlap between the two categories.**

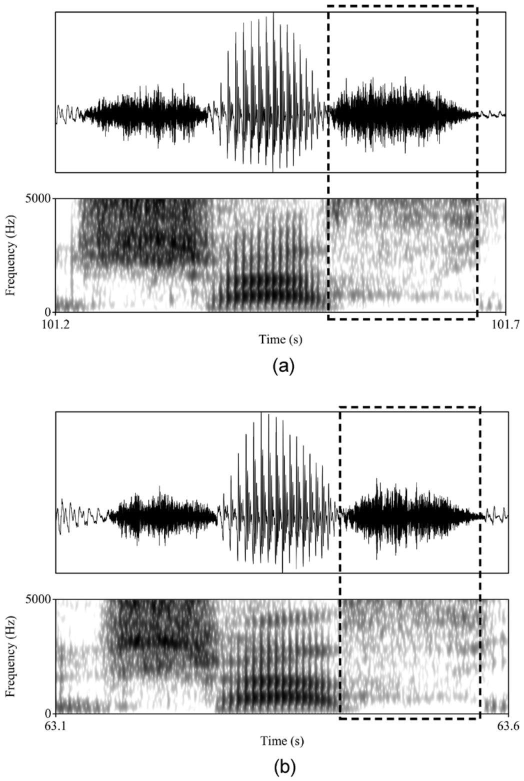


**Figure 5** **(a) A representative token of word-final voiceless fricatives (/ʃas/). The frication interval is indicated by a box with a broken line. (b) A representative token of word-final "devoiced" fricatives (/ʃaz/). The frication interval is indicated by a box with a broken line.**

without the interaction term.

For the effect of Uvoi in fricatives, we found a weak but significant effect of Uvoi for the duration of voicing during frication [$\chi^2(1)=5.5843$, $p=0.01812$]. As observed in Table 5, voicing duration during frication was slightly longer during underlyingly voiced fricatives than during underlyingly voiceless ones, although the estimated difference was 2.9 ms.

Finally, the effect of POA was not significant [$\chi^2(2)=5.6914$, $p=0.05809$]. A summary of the model for voicing duration is shown in Table 6. For all other measurements (*V1*, *C*, *F0 pre C*, *F1 pre C*), the effect of Uvoi was not significant.

## 4. Discussion

The present study addressed two questions about incomplete neutralization, using an experimental paradigm that minimized the effects of orthography and lexical factors. The first question was whether incomplete neutralization holds in non-utterance-final positions. The second question was whether stops and fricatives differ in terms of devoicing. The summary of our findings and the implications of our results are discussed in the following sections.

### 4.1 Word-final Devoicing in Utterance-medial Position

Our results show that there are significant acoustic differences between the words containing underlyingly voiced and those containing underlyingly voiceless obstruents. For stops, the duration of the preceding

**Table 5 Average values with standard deviations (in parentheses) for all measurements (fricatives). The measurement with statistical significance is indicated in boldface.**

| | Labiodental | | Alveolar | | Post alveolar | |
|---|---|---|---|---|---|---|
| | Voiceless /f/ | Devoiced /v/ | Voiceless /s/ | Devoiced /z/ | Voiceless /ʃ/ | Devoiced /ʒ/ |
| *Vowel* [ms] | 122 (22) | 125 (21) | 123 (21) | 126 (26) | 124 (20) | 130 (23) |
| ***Voicing* [ms]** | **15 (7)** | **17 (8)** | **11 (6)** | **16 (13)** | **14 (5)** | **16 (7)** |
| *Constriction* [ms] | 150 (30) | 141 (34) | 163 (29) | 155 (30) | 165 (29) | 163 (26) |
| *F0 pre C* [Hz] | 241 (72) | 243 (75) | 247 (77) | 248 (79) | 250 (74) | 248 (73) |
| *F1 pre C* [Hz] | 455 (130) | 459 (119) | 422 (110) | 423 (98) | 453 (132) | 445 (119) |

**Table 6 A model summary for voicing duration (for fricatives). Intercept is centered to "devoiced" level across POA. *p*-values smaller than 0.05 but greater than 0.01 (after being adjusted for multiple comparisons, by Bonferroni correction) are indicated with "*".**

| *Predictor* | Estimate ($\beta$) | Standard Error | t-value | |
|---|---|---|---|---|
| *Intercept ("devoiced")* | 16.3051 | 1.3362 | 12.203 | |
| *Uvoi (voiceless)* | −2.8941 | 1.129 | −2.563 | * |
| *POA (labiodental)* | −1.5465 | 0.7633 | −2.026 | n.s. |
| *(postalveolar)* | 1.3793 | 0.6238 | 2.211 | n.s. |

vowel was significantly longer before the underlyingly voiced stop than before the underlyingly voiceless stop. For fricatives, voicing duration was longer in underlyingly voiced fricatives than in underlyingly voiceless ones. The results in general confirm that incomplete neutralization is indeed observed in utterance-medial position, even when speakers produced nonce nouns in response to the auditory stimuli. Moreover, incomplete neutralization was observed even when the effects of orthography or word-specific effects are minimized.

The difference we observed in the preceding vowel duration is consistent with the results of Kulikov (2012), who examined final devoicing in stop consonants in utterance-medial position. The direction and the magnitude of the difference observed in the present study is fairly close to that in Kulikov (2012); our LME regression model estimated 6.5 ms vowel difference as a function of underlying voicing of stop, while the difference observed in Kulikov (2012) was 6 ms.

On the other hand, while Kulikov (2012) observed a difference in F1 before the final obstruent, no such differences were found in the present study. The difference may come from the experimental tasks: Kulikov examined existing words by using a reading task, while this study used the pronunciation of nonce words in response to auditory prompts. This in turn suggests that different task formats may influence how incomplete neutralization manifests itself.

Another possibility is a regional difference: Kulikov (2012, p. 27) recorded his data in Tambov, while our speakers were from Orenburg. However, this explanation is probably unlikely, as both are presumably the same variety of Russian, in that both of them have a voicing contrast and a final devoicing process (Avanesov and Orlova 1965, p. 84).

### 4.2 The Difference between Stops and Fricatives in the Implementation of Devoicing

As discussed above, for stops, we found the effect of underlying voicing in the vowel preceding stops, *a la* Kulikov (2012). Our hypothesis for fricatives was, if stops and fricatives behave identically in terms of devoicing, we would observe the same difference. Contrary to this prediction, we found a weak effect of underlying voicing on voicing duration during frication. The results suggest that stops and fricatives show different patterns of incomplete neutralization.

Why stops and fricatives behave differently raises an important theoretical question; since stops and fricatives behave differently in terms of how they are incompletely devoiced, do they possibly undergo two different "devoicing rules"? We find this unlikely for the following reason. As reviewed in Padgett (2012), Russian sonorants do not undergo obligatory devoicing, except for gradient, non-obligatory devoicing. In contrast with sonorants, our results (Figures 2(a) and 4(a)) showed that both stops and fricatives devoiced almost obligatorily, which is qualitatively different from the case of sonorants. In this sense, stops and fricatives form a natural class, and it is natural to think that the same process targets that class.

It is unlikely that speakers can control 3-ms difference of voicing for fricatives. 3 ms corresponds to a half of one cycle for speech with 167 Hz (1000/167≒6). Therefore, where this small difference comes from is a very difficult question to address. Also for stops, it should be noted that the reason vowels are longer before voiced stops is a matter of extensive debate in the first place (e.g., Kluender et al. 1988 vs. Folwer 1992); therefore, solving this question in the context of incomplete neutralization is even more complex. However, it should be noted that, as discussed in section 4.1, a 6.5-ms difference in vowel duration for stops is fairy close to previous studies carried out in Russian (6 ms in Kulikov 2012) and in other languages (8.6 ms in Roettger et al. 2014 for German; 3.5 ms in Warner et al. 2004 for Dutch). Although this difference in stops is admittedly small, the difference is larger in stops than in fricatives. This difference may potentially be attributed to the relatively denser lexical neighborhood in the stop stimuli, compared to the fricative stimuli (See Appendix). That is, it may be that the stop stimuli activate more neighbors—and importantly, non-neutralized forms thereof—which could have resulted in incomplete neutralization. However, this is mere speculation; systematically testing how the activation of neighboring words can lead to incomplete neutralization (to the extent that it can) needs to be investigated in future studies.

Ultimately, our observation that stops and fricatives behave differently may bear on the question of how to model incomplete neutralization. There are various theoretical proposals of incomplete neutralizations, including the one deploying the interactions between phonetic context and surface underspecification (Matsui 2015), phonetic paradigm uniformity with weighted constraints (Braver 2013), the difference in exemplar clouds (e.g., Ernestus and Baayen 2007), or the combination of grammar dynamics and environmental dynamics (Gafos 2006). We cannot offer conclusive remarks here, but how each theory of incomplete neutralization can handle the differences between stops and fricatives offers an interesting theoretical challenge.

### 4.3 The Role of Prosodic Conditionings on Final Devoicing and Incomplete Neutralization

To put the current findings into a broader perspective, our results are in line with previous instrumental studies in Russian (Chen 1970, Dmitrieva et al. 2010, Kharlamov 2012, 2014 Kulikov 2012). Table 7 shows a survey of the acoustic studies of word-final devoicing in Russian. As can be seen in Table 7, Kulikov (2012) and the present study show similar results; that is, they both found an effect of underlying voicing on the preceding vowel duration. On the other hand, Dmitrieva et al. (2010) and Kharlamov (2012, 2014) found no significant difference in preceding vowel duration. Instead, they found the difference in closure/frication duration and release duration of stops. This difference may come from the difference in prosodic conditions examined: while Kulikov (2012) and the present study embedded the target words in the middle of a sentence (i.e., utterance-medial position), Dmitrieva et al. (2010) and Kharlamov (2012, 2014) asked the speakers to pro-

**Table 7 Comparison with previous acoustic studies in Russian in a chronological order. All the studies except Kulikov (2012) and the current study referred to *utterance-final* position.**

| Study | Obstruent | Task | Acoustic cues | Notes |
|---|---|---|---|---|
| Chen (1970) | /p, b, t, d, k, g/ | Reading | Preceding vowel (vd＞vls) | 1 speaker, No statistical analysis |
| Pye (1989) | /p, b, t, d, k, g, f, v, s, z, ʃ, ʒ/ | Reading | Preceding vowel; Voicing (vd＞vls); Constriction (vd＜vls); Release (vd＜vls) | No statistical analysis |
| Shrager (2010) | /t, d/ | Reading | Constriction (vd＜vls, 2 ms); Release (vd＜vls, 22 ms) | |
| Dmitrieva et al. (2010) | /p, b, t, d, k, g, f, v, s, z, ʃ, ʒ/ | Reading | Constriction (vd＜vls, 16 ms); Release (vd＜vls, 16 ms) | Both for monolingual and bilingual speakers; Lexical factors controlled; Real words |
| Kulikov (2012) | /p, b, t, d, k, g/ | Reading | Preceding vowel (vd＞vls, 6 ms); F1 pre C (vd＜vls, 20 Hz) | |
| Kharlamov (2012, 2014) | /p, b, t, d, k, g, f, v, s, z, ʃ, ʒ/ | Reading; Non-reading | Constriction (vd＜vls, 7 ms); Release (vd＜vls, 6 ms) | Lexical factors controlled; Real words |
| Current study | /p, b, t, d, k, g, f, v, s, z, ʃ, ʒ/ | Non-reading | Preceding vowel (stops, vd＞vls, 6.5 ms); Voicing (fricatives, vd＞vls, 2.9 ms) | Lexical factors controlled; Nonce words |

duce the words in isolation (i.e., utterance-final position). As discussed in section 1.4, since utterance-final position undergoes phonetic lengthening (e.g., Klatt 1976, Wightman et al. 1992), this may have affected the results of Dmitrieva et al. (2010) and Kharlamov (2012, 2014). These observations suggest that the type of prosodic boundary after the word-final obstruent may affect the acoustic profile of word-final devoicing. The effects of prosodic boundary on phonetic details have been receiving an increasing body of attention in the recent literature (e.g., Kuzla and Ernestus 2011); our study shows that prosodic differences can in fact affect how segmental devoicing is implemented.

In general, prosodic domains define the domains within which phonological rules or constraints apply (Selkirk 1980 *et seq*.). They also determine the patterns of phonetic implementation, especially through domain-edge strengthening and final lengthening (Byrd et al. 2006, Wightman et al. 1992). The present study discovered another aspect in which prosodic domains are important: they may define the way in which a particular contrast is incompletely neutralized. In order to further examine this possibility, future studies should directly compare utterance-medial position with utterance-final position.

## 5. Conclusions

The present study offers additional data regarding the acoustic manifestation of Russian word-final voiceless and voiced obstruents. The results show small but significant durational differences between words with underlyingly voiceless obstruents and their devoiced counterparts, thus demonstrating a case of incomplete neutralization. The current results offer non-trivial findings about our current understanding of incomplete neutralization. First, incomplete neutralization arises, even when orthography and word-specific effects are minimized. Second, final devoicing can be incomplete, despite the fact that the target consonants in question are in utterance-medial position, where word-final devoicing is not confounded with utterance-final devoicing. Third, stops and fricatives can behave differently in terms of how they incompletely neutralize voicing contrast. Viewed from a broader perspective, by comparing the current studies with previous studies, we raised the possibility that prosodic structures can affect phonetic implementation patterns by determining how incompletely neutralized contrasts manifest themselves.

## Acknowledgements

This paper is a substantially revised version of Chapter 5 of the first author's doctoral dissertation (Matsui 2015). The earlier versions of this paper were presented at Tsukuba Phonetics and Phonology Seminar (February 2013), PAIK (May 2013), and APAP-1 (June 2013). We would like to express our gratitude to the audience for their feedback. We are also grateful to Jason Shaw, Jaye Padgett and two anonymous reviewers for valuable input and to Helen Stickney, Donna Erickson and Tim Vance for proofreading this paper. Finally, many thanks are extended to the participants of our experiment and Liudmila V. Dokashenko for arranging the research visit to Russia. This research was partially supported by the MEXT/KAKENHI Grants-in-Aid for Scientific Research (#246727, #15H06832, #17J08493) awarded to the first author.

### Notes

1) Final devoicing has also been used as an example for positional neutralization in well-known introductory phonology textbooks (e.g., Kenstowicz 1994, Odden 2005).

2) This is not to say, however, that incomplete neutralization is impossible to deal with in formal phonological theories. See Braver (2013) for a review of various approaches to incomplete neutralization within formal phonological theories.

3) While the focus of the present study is acoustic aspects of word-final devoicing, there are a few studies on the perception and aerodynamics of incomplete neutralization in Russian (Kanibolotskaja 2009, Kharlamov 2012, 2014, Matsui 2011, 2015). Kanibolotskaja (2009) found that the underlyingly voiced and voiceless bilabial stops differ in intraoral air pressure. Matsui (2011) demonstrated that "incompletely" neutralized velar stops were perceptually distinguishable by native listeners of Russian. Kharlamov (2012, 2015) and Matsui (2015) extended Matsui's (2011) observation with a large data set.

4) We are grateful to an anonymous reviewer for suggesting this possibility.

5) $d'$ is a sensitivity measure in Signal Detection Theory (Macmillan and Creelman 2005). Higher $d'$ indicates higher sensitivity to stimuli, and $d'=0$ indicates no sensitivity to stimuli (i.e., chance-level accuracy). $d'$-values higher than 4 are near-ceiling; when unbiased, 95% correct responses would result in $d'$-value of 3.96. Those listeners who checked the stimuli did not participate in the current production experiment.

6) The current experiment does not tease apart phrase-final position from word-final position. Since the target

items were followed by a syntactic VP boundary, it may be possible that there was a phonological phrase boundary after the target words as well. See Selkirk (2000) and Truckenbrodt (1995) for overviews of the languages which place a phrase boundary before a syntactic VP boundary (although there is of course no guarantee that a VP boundary coincides with a phrase boundary in Russian). Teasing apart word-final positions from phrase-final positions is an important task for future experiments.

7) The default of lmer function uses a dummy coding (i.e., treatment contrast), in which one specific level within a factor (e.g., voiceless bilabial stops, but not voiceless stops in general) is set as a reference level (i.e., intercept). However, what we are interested in is the effect of underlying voicing *across* three different places of articulation, not the difference between one specific level (e.g., voiceless bilabial stop) with all other levels. For this reason, we manually sum-coded POA.

## References


Akishina, A. A. and S. A. Baranovskaja (2010) *Russkaja fonetika na fone obschej* [Russian phonetics in a general phonetic perspective]. 5th edition. Moscow: LKI.

Avanesov, I. R. and V. G. Orlova (1965) *Russkaja dialectologija* [Russian dialectology]. Moskva: Nauka.

Barr, D. J., R. Levy, C. Scheepers and H. J. Tily (2013) "Random effects structure for confirmatory hypothesis testing: Keep it maximal." *Journal of Memory and Language* 68, 255–278.

Baese-Berk, M. and M. Goldrick (2009) "Mechanisms of interaction in speech production." *Language and Cognitive Processes* 24, 527–554.

Barnes, J. (2006) *Strength and weakness at the interface: Positional neutralization in phonetics and phonology*. Berlin: Mouton de Gruyter.

Bates, D., M. Maechler, B. Bolker and S. Walker (2014) Lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7.

Blevins, J. (2006) "A theoretical synopsis of Evolutionary Phonology." *Theoretical Linguistics* 32, 117–165.

Boersma, P. and D. Weenink (2010) Praat: Doing phonetics by computer (Version 5.1.31). http://www.praat.org/

Braver, A. (2013) *Degrees of incompletenesss in neutralization: Paradigm uniformity in a phonetics with weighted constraints*. Doctoral dissertation, Rutgers University.

Byrd, D., J. Krivokapić and S. Lee (2006) "How far, how long: On the temporal scope of phrase boundary effects." *Journal of the Acoustical Society of America* 120, 1589–1599.

Charles-Luce, J. (1985) "Word-final devoicing in German: Effects of phonetic and sentential contexts." *Journal of Phonetics* 13, 309–324.

Chen, M. (1970) "Vowel length variation as a function of the voicing of the consonant environment." *Phonetica* 22, 129–139.

Cubberley, P. (2002) *Russian: A linguistic introduction*. Cambridge, UK: Cambridge University Press.

Davison, L. (2016) "Variability in the implementation of voicing in American English obstruents." *Journal of Phonetics* 54, 35–50.

Dinnsen, D. A. and J. Charles-Luce (1984) "Phonological neutralization, phonetic implementation and individual differences." *Journal of Phonetics* 12, 49–60.

Dmitrieva, O., A. Jongman and J. Sereno (2010) "Phonological neutralization by native and non-native speakers: The case of Russian final devoicing." *Journal of Phonetics* 38, 483–492.

Ernestus, M. and H. R. Baayen (2006) "The functionality of incomplete neutralization in Dutch: The case of past tense formation." *Laboratory Phonology* 8, 27–49.

Ernestus, M. and H. R. Baayen (2007) "Intraparadigmatic effects on the perception of voice." In E. van der Torre and J. van de Weijer (eds.) *Voicing in Dutch*, 153–173. Amsterdam: Benjamins.

Fourakis, M. and G. Iverson (1984) "On the 'incomplete neutralization' of German final obstruents." *Phonetica* 41, 140–149.

Fourakis, M. and R. Port (1986) "Stop Epenthesis in English." *Journal of Phonetics* 14(2), 197–221.

Fowler, C. A. (1992) "Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here." *Journal of Phonetics* 20, 143–165.

Gafos, A. (2006) "Dynamics in grammar: Comment on Ladd and Ernestus and Baayen." *Laboratory Phonology* 8, 51–79.

Hall, K. C., B. Allen, M. Fry, S. Mackie and M. McAuliffe. (2016) Phonological CorpusTools, Version 1.2.0. https://github.com/PhonologicalCorpusTools/CorpusTools/releases

Halle, M. (1959) *The sound pattern of Russian*. Hague: Mouton and Co., Printers.

Hamilton, W. (1980) *Introduction to Russian phonology and word structure*. Ohio: Slavica Publishers.

Herd, W., A. Jongman and J. Sereno (2010) "An acoustic and perceptual analysis of /t/ and /d/ flaps in American English." *Journal of Phonetics* 38(4), 504–516.

Jassem, W. and L. Richter (1989) "Neutralization of voicing in Polish obstruents." *Journal of Phonetics* 17, 317–325.

Kanibolotskaja, S. (2009) *Aerodynamic study of word-final devoicing in Russian*. MA Forum paper, University of Toronto.

Kenstowicz, M. (1994) *Phonology in generative grammar*. Malden: Blackwell.

Kharlamov, V. (2012) *Incomplete neutralization and task effects in experimentally-elicited speech: Evidence from the production and perception of word-final devoicing in Russian*. Doctoral dissertation, University of Ottawa.

Kharlamov, V. (2014) "Incomplete neutralization of the voicing contrast in word-final obstruents in Russian: Phonological, lexical, and methodological influences." *Journal of Phonetics* 43, 47–56.

Kharlamov, V. (2015) “Perception of incompletely neutralized voicing cues in word-final obstruents: The role of differences in production context.” *Journal of the Association for Laboratory Phonology* 6 (2), 147–165.

Kingston, J. and R. Diehl (1994) “Phonetic knowledge.” *Language* 70, 419–454.

Klatt, D. H. (1976) “Linguistic use of segmental duration in English: Acoustic and perceptual evidence.” *Journal of the Acoustical Society of America* 59(5), 1208–1221.

Kluender, K. and R. Diehl and B. Wright (1988) “Vowel-length differences before voiced and voiceless consonants: An auditory explanation.” *Journal of Phonetics* 16, 153–169.

Kulikov, V. (2012) *Voicing and voice assimilation in Russian stops*. Doctoral dissertation, The University of Iowa.

Kuzla, C. and M. Ernestus (2011) “Prosodic conditioning of phonetic detail in German plosives.” *Journal of Phonetics* 39(2), 143–155.

Ladefoged, P. and I. Maddieson (1996) *The sounds of the world’s languages.* New Jersey: Wiley-Blackwell.

Lyashevskaya, O. N. and S. A. Sharov (2009) *Chastotnyj slovar’ sovremennogo russkogo jazyka (na materialax Natsional’nogo korpusa russkogo jazyka)* [Frequency dictionary of Russian language (in the materials of the Russian National Corpus)]. Moscow: Azbukovnik. http://dict.ruslang.ru/freq.php?

Macmillan, N. A. and D. C. Creelman (2005) *Detection theory: A user’s guide*. 2nd edition. Mahwah: Lawrence Erlbaum Associates Publishers.

Matsui, M. (2011) “The identifiability and discriminability between incompletely neutralized sounds: Evidence from Russian.” *Proceedings of the 17th International Congress of Phonetic Sciences*, *Hong Kong*, *Hong Kong Convention and Exhibition Centre (ICPhS 17)*, 1342–1345.

Matsui, M. (2015) *Roshiago ni okeru yūseisei no tairitsu to tairitsu no jakka: Onkyō to chikaku.* [Voicing contrast and contrast reduction in Russian: Acoustics and perception]. Doctoral dissertation, Hiroshima University.

Munson, B. (2007) “Lexical access, lexical representation and vowel articulation.” *Laboratory Phonology* 9, 201–228.

Munson, B. and N. P. Solomon (2004) “The effect of phonological neighbourhood density on vowel articulation.” *Journal of Speech*, *Language*, *and Hearing Research* 47, 1048–1058.

Myers, S. (2012) “Final devoicing: Production and perception studies.” In T. Borowsky, S. Kawahara, T. Shinya and M. Sugahara (eds.) *Prosody matters: Essays in honor of Elisabeth Selkirk*, 148–180. London: Equinox Press.

Myers, S. and J. Padgett (2014) “Domain generalization in artificial language learning.” *Phonology* 31, 399–433.

Ohala, J. (1983) “The origin of sound patterns in vocal tract constraints.” In P. F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer-Verlag.

Padgett, J. (2012) “The role of prosody in Russian voicing.” In T. Borowsky, S. Kawahara, T. Shinya and M. Sugahara (eds.) *Prosody matters: Essays in honor of Lisa Selkirk*, 181–207. London: Equinox Press.

Pierrehumbert, J. (2002) “Word-specific phonetics.” *Laboratory Phonology* 7, 101–139.

Piroth, H. G. and P. M. Janker (2004) “Speaker-dependent differences in voicing and devoicing of German obstruents.” *Journal of Phonetics* 32, 81–109.

Port, R. F. and M. O’Dell (1985) “Neutralization of syllable-final devoicing in German.” *Journal of Phonetics* 13, 455–471.

Port, R. and P. Crawford (1989) “Incomplete neutralization and pragmatics in German.” *Journal of Phonetics* 17, 257–282.

Port, R. and A. P. Leary (2005) “Against formal phonology.” *Language* 81, 927–964.

Pye, S. (1986) “Word-final devoicing of obstruents in Russian.” *Cambridge Papers in Ponetics and Experimental Linguistics* 5, 1–10.

R Core Team (2014) R: A language and environment for statistical computing (Version 3.1.0). R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Roettger, T. B., B. Winter and S. Grawunder (2011) “The robustness of incomplete neutralization in German.” *Proceedings of the 17th International Congress of Phonetic Sciences*, *Hong Kong*, *Hong Kong Convention and Exhibition Centre (ICPhS 17)*, 1722–1725.

Roettger, T. B., B. Winter and S. Grawunder (2014) “Assessing incomplete neutralization of final devoicing in German.” *Journal of Phonetics* 43, 11–25.

Selkirk, E. (1980) “The role of prosodic categories in English word stress.” *Linguistic Inquiry* 11, 563–605.

Selkirk, E. (2000) “The interaction of constraints in prosodic phrasing.” In M. Horne (ed.), *Prosody: Theory and experiments*, 231–261. Dordrecht: Kluwer.

Shrager, M. (2012) “Neutralization of word-final voicing in Russian.” *Journal of Slavic Linguistics* 20(1), 71–99.

Smith, C. (1997) “The devoicing of /z/ in American English: Effects of local and prosodic context.” *Journal of Phonetics* 25, 471–500.

Smith, C. (2003) “Vowel devoicing in contemporary French.” *French Language Studies* 13, 177–194.

Solé, M. J. (2002) “Assimilatory processes and aerodynamic factors.” In C. Gussenhoven and N. Warner (eds.) *Papers in laboratory phonology* 7, 351–386. Berlin: Mouton de Gruyter.

Truckenbrodt, H. (1995) *Phonological phrases: Their relation to syntax, focus, and prominence*. Ph.D. dissertation, MIT.

VanDam, M. and R. F. Port (2005) VOT is shorter in high frequency words. Poster presented at the 149th Meeting of the Acoustical Society of America, Vancouver, BC, Canada.

Wade, T. (2010) *A comprehensive Russian grammar*. 3rd edition. Malden: Wiley-Blackwell.

Warner, N. and A. Jongman, J. A. Sereno and R. Kemps (2004) “Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch." *Journal of Phonetics* 32, 251–276.

Warner, N., E. Good, A. Jongman and J. A. Sereno (2006) "Orthographic vs. morphological incomplete neutralization effects." *Journal of Phonetics* 34, 285–293.

Wightman, C. W., S. Shattuck-Hufnagel, M. Ostendorf and P. J. Price (1992) "Segmental durations in the vicinity of prosodic boundaries." *Journal of the Acoustical Society of America* 91, 1707–1717.

Wright, R. A. (2004) "Factors of lexical competition in vowel articulation." In J. Local, R. Ogden and R. Temple (eds.) *Papers in laboratory phonology* 6, 75–87. Cambridge, UK: Cambridge University Press.

Yanushevskaya, I. and D. Bunčić (2015) "Illustrations of the IPA: Russian." *Journal of the International Phonetic Association* 45(2), 221–228.

Yu, A. C. L. (2007) "Tonal phonetic analogy." *Proceedings of the 16th International Congress of Phonetic Sciences*, *Saarbrücken*, *Saarland University (ICPhS 16)*, 1749–1752.

Żygis, M., S. Fuchs and L. L. König (2012) "Phonetic explanations for the infrequency of voiced sibilant affricates across languages." *Journal of the Association for Laboratory Phonology* 3, 299–336.

(Received Nov.13, 2016, Accepted Jul. 10, 2017)

**Appendix Speech Materials (shown in phonemic representation) with substitution neighborhood density in parentheses. See section 2.2 for further details of the neighborhood calculation.**

Stops

| Bilabial | | Dental | | Velar | |
|---|---|---|---|---|---|
| Voiceless | Voiced | Voiceless | Voiced | Voiceless | Voiced |
| /rʲip/ (7) | /rʲib/ (3) | /zʲit/ (8) | /zʲid/ (8) | /prʲik/ (10) | /prʲig/ (5) |
| /lʲep/ (9) | /lʲeb/ (8) | /ʒet/ (6) | /ʒed/ (5) | /dnʲek/ (0) | /dnʲeg/ (1) |
| /tʃap/ (9) | /tʃab/ (9) | /xrat/ (6) | /xrad/ (3) | /tsak/ (9) | /tsag/ (4) |
| /nop/ (8) | /nob/ (7) | /tsot/ (10) | /tsod/ (8) | /zok/ (11) | /zog/ (8) |
| /mup/ (5) | /mub/ (6) | /ʃtut/ (1) | /ʃtud/ (0) | /bruk/ (2) | /brug/ (5) |

Fricatives

| Labiodental | | Alveolar | | Post alveolar | |
|---|---|---|---|---|---|
| Voiceless | Voiced | Voiceless | Voiced | Voiceless | Voiced |
| /krʲif/ (3) | /krʲiv/ (3) | /zʲis/ (3) | /zʲiz/ (6) | /prʲiʃ/ (3) | /prʲiʒ/ (3) |
| /tsef/ (3) | /tsev/ (5) | /sʲes/ (9) | /sʲez/ (6) | /nʲeʃ/ (4) | /nʲeʒ/ (5) |
| /naf/ (4) | /nav/ (5) | /ʃas/ (8) | /ʃaz/ (8) | /plaʃ/ (7) | /plaʒ/ (7) |
| /ʒof/ (0) | /ʒov/ (5) | /dros/ (4) | /droz/ (2) | /moʃ/ (6) | /moʒ/ (9) |
| /ruf/ (4) | /ruv/ (4) | /plus/ (6) | /pluz/ (2) | /luʃ/ (7) | /luʒ/ (5) |