

Acquisition of the Feature [+Voice] in Word-Final Position by Native Speakers of Russian

Ala Simonchyk

Indiana University

Abstract

The current study is a pilot investigation that set out to observe the acquisition of voicing contrasts word-finally by late learners of English who were native speakers of Russian. English obstruents preserve the feature [+voice] word-finally, whereas Russian loses it due to the phonological rule of word-final devoicing. The study sought to answer whether advanced Russian learners of English could acquire the feature [+voice] word-finally, and if they did, what acoustic cues they used to encode voicing. The participants of the study were four advanced Russian learners of English and four American native speakers of English who served as a comparison group. None of the American English speakers had any prior knowledge of a neutralizing language. The participants performed a picture-naming task. Minimal pairs were excluded to avoid the effect of hyperarticulation. The results of the study suggested that advanced Russian learners of English were able to acquire the feature [+voice] word-finally if English had become their dominant language in everyday life. However, unlike the American participants, the Russian learners of English did not manipulate vowel durations with respect to the voicing status of the following consonants. Instead, the Russian learners used durations of closure and voicing into closure to encode voicing. The current study contributes to the understanding of the acquisition of marked phonological processes, as well as the effect of extralinguistic factors on phonological competence.

Keywords: neutralization, word-final devoicing, markedness, voicing cues, Russian, acquisition of [+voice]

1. Introduction


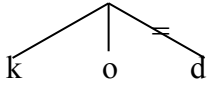
Voicing neutralization has already become a classic research area in the field of linguistics and second language studies. It is not surprising since a large number of languages exhibit word-final devoicing, i.e. the underlying voicing contrast in obstruents does not surface phonetically at the end of words. For instance, in Russian, /kod/ ‘code’ and /kot/ ‘cat’ are both realized as [kot]. Neutralizing languages include almost all Slavic languages, as well as German, Dutch, Catalan, Korean, Mongolian, Turkish, Afrikaans, amongst others. In nonneutralizing languages, such as English, the voicing contrast is realized word-finally, e.g., /kod/ ‘code’ and /kot/ ‘coat’ are pronounced as [kod] and [kot] respectively.

According to the Markedness Differential Hypothesis (Eckman, 1977), areas in a target language that are different from the native language and more marked cross-linguistically are harder to acquire. Voiced consonants are considered more marked than voiceless consonants because voiced consonants occur less frequently in the world’s languages. Moreover, voicing contrasts word-finally are more marked in comparison to voicing contrasts word-initially or word-medially. Therefore, learners whose native language neutralizes voicing contrasts in the word-final position, for example, Russian, have more difficulty acquiring a language which retains this contrast, such as English, than vice versa, i.e. when an English speaker is learning Russian. In other words, Russian learners of English have to acquire something ‘new’, i.e. a feature [+voice] for final obstruents, whereas English learners of Russian have to learn to stop using or suppress the feature [+voice] in that specific environment. From the perspective of autosegmental phonology (Goldsmith, 1990), Russian learners of English have to spread the laryngeal

node of voice to the final consonant, whereas English learners of Russian have to delink it (Table 1).

Table 1

Autosegmental perspective on the acquisition of word-final voicing and devoicing

Target	Word-final voicing in English	Word-final devoicing in Russian
Learners	Russian learners of English	English learners of Russian
Linguistic process	<p>Spreading [+voice] [-voice]</p> 	<p>Delinking [+voice]</p> 
Word 'code'	English: /kod/ [kod]	Russian: /kod/ [kot]

The goal of the current study is to investigate whether advanced late learners of English with a first language of Russian can acquire word-final voicing. In order to ensure that participants of the study have achieved a high level of English proficiency, two criteria have been adopted. First, only learners who have been enrolled in a graduate program at an American university for at least one year will be recruited for the study. It is assumed that if the participants are able to enter an American university and maintain their graduate status for at least two semesters successfully, their level of English proficiency is relatively high. Second, the participants must have spent a minimum of two years in the United States. Flege and Fletcher (1992) have found that length of residence has the greatest effect on the degree of foreign accent during the first year of stay in a second language environment. For highly experienced learners additional years spent in the country of the target language are unlikely to bring significant changes in the degree of foreign accent. Since during the first year the participants might still be

undergoing developmental changes, it was decided to exclude the first couple of years in order to test phonological grammars that have reached a plateau in this respect.

The study targets not only advanced learners of English but also late learners. For the purposes of this project, late learners are considered to be learners who arrived in the United States after the age of 18. Most of them have already received formal instruction in their native country prior to their arrival in the USA. However, the phonological development of these learners is not very high because most teachers of English are native speakers of Russian with an accented phonology and the overall amount of native input that Russian learners of English receive in Russia is not very substantial. The age of 18 is used as a cut-off point because the earlier learners are immersed into the second language environment, the higher the chance that their pronunciation will be nativelike. The proponents of the critical period hypothesis believe that after a specific age successful learning cannot take place, especially in the acquisition of phonology. Long (1990) considers the age of six to be such a point, i.e. languages that are learned later are likely to be accented. Scovel (1988) believes that the threshold occurs at the age of 12, whereas Patkowski (1990) mentions the age of 15. However, in a study by Flege, Munro, and Mackay (1995), a few Italian native speakers who had started learning English after the age of 15 received nativelike ratings. In another study by Flege, Frieda, and Nozawa (1997), Italian speakers who rarely used their native language received nativelike ratings in English even if they started learning English at the age of 16. Recruiting learners who are over 18 helps eliminate advantages that onset at an early age can have on pronunciation. However, there is evidence that if late learners who are 18 and over are

exposed to native input, as well as receive pronunciation training, their ratings can be similar to those of native speakers (Bongaerts, Summeren, Planken & Schils, 1997).

Thus, the current study seeks to establish whether native speakers of a neutralizing language who are proficient late learners of a nonneutralizing language can attain the word-final voicing distinction.

2. Previous Research

Experimental studies on voicing neutralization have revealed a set of parameters or acoustic cues that can be affected by the feature [+voice]. They are: (a) duration of the preceding vowel; (b) duration of voicing into closure for stops or frication for fricatives; (c) duration of closure for stops or frication for fricatives; (d) duration of release for stops only.

Not all phonetic cues are employed simultaneously to maintain the distinction between voiced and voiceless consonants. Overall, there is considerable variation of results even in studies done on the same language. Port and O'Dell (1985) investigated voicing neutralization in German stops using a reading task. They found that vowels were significantly longer before underlying voiced stops than they were before voiceless stops and voicing into closure was significantly longer for voiced stops than for voiceless stops. The strongest effect was found for the duration of the release. In a follow-up study, Port and Crawford (1989) replicated these results. Using discriminant analysis, they concluded that voicing was not a property of a single segment but was distributed across the entire word. Differences in vowel duration, closure duration and release all contributed to the differentiation of voiced consonants from voiceless. The participants of the study did not use just one acoustic cue but a number of weak cues to express the voicing contrast.

However, Fourakis and Iverson (1984), as well as Piroth and Janker (2004), did not find any consistent differences in vowel duration or voicing into closure for voiced and voiceless consonants in German. Warner, Jongman, Sereno, and Kemps (2004) examined Dutch. The results showed that the main acoustic cues employed by the speakers to distinguish between voiced and voiceless consonants were vowel duration and release duration. Dinnsen and Charles-Luce (1984) investigated voicing neutralization in Catalan. The results of the study showed that vowels preceding target consonants were shorter before voiceless than before voiced consonants and closure duration was longer in voiceless consonants than in voiced. They also found that individual speakers employed different strategies to distinguish voiced and voiceless consonants in their production. Some speakers manipulated the duration of the preceding vowel, whereas others changed the closure duration of the target consonant. The findings of the study conducted by Slowiaczek and Dinnsen (1985) suggested that in Polish vowel duration was the main acoustic cue used by the target consonants. Voicing into closure was employed only for labial consonants. However, another study on Polish by Jassem and Richter (1989) found that although some speakers did use vowel duration and closure duration as acoustic cues, the differences were not significant.

Experimental studies on Russian provide controversial results with respect to which acoustic cues are relevant in distinguishing voiced and voiceless consonants. The earliest work done by Chen (1970) investigated whether Russian speakers used vowel duration as an acoustic cue to encode voicing. The results of the study showed that, indeed, Russian speakers produced longer vowels before voiced stops than before voiceless stops. Later work by Pye (1986) confirmed that vowel duration was used as an

acoustic cue to differentiate voiced and voiceless consonants but also identified two additional acoustic cues: consonant duration and voicing into closure. However, statistical analyses were not performed on the results to identify whether the differences in acoustic cues were significant. A follow-up study was conducted by Shrager (2012) to examine voicing neutralization of coronal stops in Russian. The results showed that the most salient acoustic cue was the release. Differences in vowel duration and voicing into closure were found to be insignificant. A relatively recent study by Dmitrieva, Jongman, and Sereno (2010) did not support the claim that vowel duration was used as an acoustic cue to encode voicing in Russian. Their results suggested that Russian native speakers distinguished final obstruents that contrast in underlying voicing by two measures: duration of the closure or frication and duration of the release. Kharlamov (2012) also found that vowel differences before voiced and voiceless consonants were not significant in Russian. The results were significant only for closure or friction duration, voicing into closure or frication and release.

Why is there such reported variation in acoustic cues? Acoustic cues depend on the degree of voicing in the output. When voicing fluctuates, acoustic cues start to vary as well. The degree of voicing in production can depend on a number of factors. For example, task effects. If participants are asked to read a list of minimal pairs, they can unintentionally, or under certain circumstances intentionally, hyperarticulate their production of voiced consonants in order to distinguish them from their voiceless counterparts. On the other hand, if participants are recorded spontaneously saying target words, the voicing distinction is neutralized, e.g., a study by Fourakis and Iverson (1984). Another potential confounding factor is the presence of orthography. It can be argued that

orthography triggers voicing in the production of voiced consonants word-finally in Russian because it spells out underlying representations.

Port and Crawford (1989) conducted a set of experiments to identify whether the degree of voicing depends on such factors as presence vs. absence of orthography, presence vs. absence of attention to the voicing contrasts (or minimal pairs). The results showed that orthography did not play a decisive role in the degree of voicing but the speakers indeed could manipulate the degree of voicing when producing a contrast between voiced and voiceless consonants for pragmatic reasons, i.e. to make the distinction between the words in a minimal pair more salient for the listener. Also, reading a word list preserved more voicing than oral production.

Kharlamov (2012) also examined task-dependent influences of orthography and minimal pairs on voicing neutralization in Russian. The results of the study showed that differences in consonantal durations were found independently of the presence of orthographic representations or minimal pairs. However, both orthography and minimal pairs had a biasing effect on the production of voicing with the lexical effect being more prominent than the effect of orthography.

Therefore, it is not surprising that the acoustic cues reported in a number of studies differed in their significance. However, it is important to mention that irrespective of the language, four basic acoustic cues that are used to distinguish voiced and voiceless consonants behave in a similar manner: voiced consonants have longer preceding vowels and durations of voicing into closure or frication, whereas voiceless consonants have longer durations of closure or frication as well as release for stops. Different languages seem to have different primary acoustic cues, although the data are not always very

straightforward. For example, in Russian, it is still not clear whether vowel durations change depending on the voicing status of the subsequent consonants. However, in English, although it is not a neutralizing language, vowel duration is the most important acoustic cue that distinguishes voiced and voiceless consonants. English vowels before voiceless consonants are 54% as long as vowels before voiced consonants (Mack, 1982).

Smith, Hayes-Harb, Bruss, and Harker (2009) conducted a study to investigate how well German learners of English could produce voiced obstruents word-finally in English while continuing to neutralize the voicing contrast in German. They used similar word pairs in the two languages (e.g., English 'rod/rot' vs. German 'Rad/Rat') embedded in a carrier phrase that participants had to read. The results of the study demonstrated that German learners of English were able to distinguish voiced and voiceless word-final stops in English in their production but their voicing contrast was less robust than that of English native speakers. For example, vowel durations produced by German learners were very similar to those produced by English native speakers but closure durations and durations of voicing into closure were significantly different.

Simon (2010) examined the production of final obstruents and voicing assimilation across word boundaries by Dutch learners of English in natural conversations. The results of the study showed that the participants transferred the word-final devoicing rule from Dutch to English. Word-final voiced fricatives were devoiced much more than voiced stops. Previous explicit instruction on word-final devoicing but not cross-word voicing assimilation resulted in less transfer of the former rule but not the latter. The presence or absence of voicing was coded on the basis of auditory judgments; therefore, no information on primary acoustic cues was available.

Neither Smith et al. (2009) nor Simon (2010) mentioned the English proficiency levels of their subjects explicitly. Smith et al. provided details from their participants' profiles, such as length of residence (nine months to 17 years), age of arrival (16 - 26 years of age) and length of formal instruction (7 - 14 years). Eleven out of 13 participants reported that they were fluent in English. The participants of Simon's study also reported that they were fluent in English and allegedly had "an advanced pronunciation" (p. 68). There was no information provided about the participants' experience abroad or years of formal instruction. Overall, it can be assumed that both studies recruited advanced learners of English. Most probably, the participants in Simon's study had no experience living in an English-speaking country, whereas participants in the study by Smith et al. were living in the United States during the time of data collection.

The rule of word-final voicing seems to pose challenges for learners of neutralizing languages. Although learners are able to distinguish voiced and voiceless obstruents word-finally in their production, they use different acoustic cues that are not as robust as those of native speakers. The proposed study will try to answer the following questions:

1. Can advanced late learners of English who are native speakers of Russian acquire the feature [+voice] word-finally in English?
2. What acoustic cues do Russian learners of English employ to distinguish voiced and voiceless consonants word-finally in English?

It can be hypothesized that proficient Russian learners of English who have been exposed to English living in the second language environment can produce voicing contrasts word-finally similarly to native speakers of American English. Learners

continuously hear word-final voicing contrasts in the input they receive. Although word-final obstruents themselves are not phonetically salient, the lexical load they carry can increase their saliency and importance. If Russian learners neutralize the voicing distinction word-finally in their production and say, for instance, [bʌk] instead of [bʌg] in a sentence “I see a bug!”, they can mislead their listeners into believing that they see a one-dollar bill instead of an insect. Thus, applying the neutralization rule from Russian can result in a communication breakdown, which can raise learners’ awareness of word-final voiced obstruents in English. Moreover, the feature [+voice] is present in the phonemic inventory of Russian, which should make it easier for acquisition. Also, due to the high level of English proficiency of the Russian participants in the current study, they can process and produce language relatively quickly and automatically, thereby allowing them to pay more attention to their pronunciation.

With respect to the second research question, it is hypothesized that similar to American native speakers, Russian learners of English will use vowel duration as the main acoustic cue to distinguish voiced and voiceless consonants word-finally. Taking into account the fact that there is evidence that vowel durations depend on the voicing of the following consonants in native Russian (Chen, 1970; Pye, 1986), it is expected that the same strategy will be utilized in producing word-final voiced obstruents in English, especially considering that vowel duration is the most important voicing cue in English.

To date, no study has empirically investigated the acquisition of word-final voiced obstruents in L2 English by Russian native speakers. Thorough screening of participants ensured that only late learners with high English proficiency and substantial experience living in the United States were recruited for the study. Furthermore, an effort was made

to elicit productions of word-final obstruents that are not characterized by hyperarticulation. In order to avoid task effects, such as the influence of orthography and minimal pairs, the participants of the study performed a picture-naming task with multiple distractors that aimed to divert their attention from the target phenomenon. The next section offers further details on the methods used in the current study.

3. Method

3.1. Participants

The participants of the pilot study were four Russian learners of English (2 females, 2 males) and four native speakers of American English (2 females, 2 males) who served as a comparison group. The native speakers of American English (age 23 - 29) were graduate students at a large American university in the US with no prior knowledge of Russian or another neutralizing language, such as German, Catalan, Turkish or any Slavic language. The Russian learners of English (age 28 - 31) were originally from Russia and arrived in the United States at the age of 21 - 23 (Table 2).

Table 2

Russian learners' profiles

Participant	1	2	3	4
Gender	F	F	M	M
Age	28	31	28	31
English proficiency (self-report)	nativelike	nativelike	nativelike	advanced
Age of initial English instruction	7	10	11	10
Age of arrival in the US	23	22	23	21
Length of residence in the US (years)	5	10	6	8
Daily amount of L1 use (%)	10	3	5	70
Daily amount of L2 use (%)	90	97	95	30

The Russian learners had formal English instruction in Russia before coming to the United States. Their length of residence in the United States was between 5 - 10

years. The Russian participants were highly proficient learners of English enrolled as PhD students in the same American university in the US. Three of the four participants reported using English much more often than Russian in their daily life.

3.2. Materials

Eighteen monosyllabic target words were selected for the study: ‘sheep’, ‘whip’, ‘loop’, ‘rib’, ‘cube’, ‘web’, ‘beet’, ‘foot’, ‘net’, ‘seed’, ‘hood’, ‘bed’, ‘cheek’, ‘book’, ‘neck’, ‘pig’, ‘fig’, ‘leg’. In order to control for intervening variables, such as phonetic environment, durational and frequency effects, a number of criteria were employed. The target words were singular concrete nouns that could be represented with a picture. They were lexically frequent as checked in the Corpus of Contemporary American English (Davies, 2008). The word ‘beet’ had the lowest frequency (720 per million). The syllable structure of the target words was CVC. The word-final consonants were stops equally distributed according to place of articulation: six labials (p / b), six alveolars (t / d) and six velars (k / g). The preceding vowels were all [-low], such as /i, ɪ, u, ʊ, ε/. Low vowels were excluded because they are on average 20 ms longer than high and mid vowels due to the longer movement of the jaw (Reetz & Jongman, 2009). I decided not to include /e/, /o/ because Russian participants often produce them as diphthongs /eɪ/ and /oʊ/ respectively. Five vowels /i, ɪ, u, ʊ, ε/ were chosen as this was the minimal number that could guarantee a set of words that would satisfy the other requirements. No minimal pairs were used that alternated in the target consonants. This was done to avoid attracting participants’ attention to the phenomenon.

Twenty-four distractors were added to the list: ‘can’, ‘pen’, ‘fan’, ‘phone’, ‘bun’, ‘ring’, ‘wing’, ‘room’, ‘team’, ‘bee’, ‘key’, ‘door’, ‘car’, ‘shoe’, ‘rose’, ‘nose’, ‘leaf’,

‘fish’, ‘bush’, ‘ball’, ‘doll’, ‘bell’, ‘pill’, ‘wheel’. All distractors were semantically associated either with the target words or with other distractors, for example, body parts (‘leg’, ‘foot’, ‘cheek’, ‘nose’), animals (‘sheep’, ‘pig’, ‘fish’, ‘bee’), food (‘fig’, ‘rib’, ‘beet’, ‘seed’, ‘bun’), round objects (‘loop’, ‘web’, ‘pill’, ‘wheel’, ‘ring’) etc. The distractors were monosyllabic words of CVC and CV structure that ended either in a nasal, vowel or fricative. They were also lexically frequent items as checked in the Corpus of Contemporary American English (Davies, 2008).

Thus, there were 42 items total, which were divided into three blocks (ABC) with six target items and eight distractors per block. Six target items in each block represented six target consonants. All items were repeated in two cycles, which resulted in the following order of blocks: ABC + BCA. The items in each block were pseudorandomized for each cycle. Target items with the same place of articulation for the final consonant, e.g., ‘beet’ and ‘seed’ were never placed next to each other.

Each word was matched to a colorful picture on a white background. The pictures were tested on two people prior to data collection to see how effective they were at eliciting the target words. Some of the pictures were replaced after the trial, e.g., the picture of a webpage to elicit the word ‘web’ was replaced with a picture of a spider-web. The benefit of using pictures was that it helped reduce the effects of orthography. The word-final voicing distinction that is reflected in English spelling could have alerted Russian speakers to produce more voicing during the elicitation stage than they would have normally produced in their natural speech. Thus, using pictures allowed for elicitation that better approximated the encoded phonological forms of the target words than reading a wordlist.

3.3. Procedure

The participants were tested individually by a researcher who was a native speaker of Russian in a sound-proof recording studio to ensure high quality of the recordings. The procedure included two stages: practice stage and elicitation stage. In order for the Russian learners to be in their English mode, English was used throughout the session. The goal of the practice stage was to familiarize participants with the pictures and ensure that they would produce target words during an elicitation stage. For example, a participant was shown a flashcard with a question “What is it?” on the top, a picture of an object, e.g., a bun, and a sentence “It’s a b___.” The first letter of the word was provided. The participant was asked to guess the word. If the participant answered “bread”, the researcher would say, “Can you think of another word?” and would allow two more attempts. After three mistakes the researcher would provide the correct answer. The pictures that caused difficulty, e.g., a wrong guess or a long pause before answering, were put aside. When all the pictures had been presented, the “problematic” ones were practiced one more time. Then the participant was asked whether all the pictures were clear and whether he or she was ready to be recorded. The practice stage was not recorded and lasted about four minutes.

During the elicitation stage, the participants performed a picture-naming task using a PowerPoint presentation. The participants saw exactly the same pictures from the practice stage and had four seconds to produce the sentence “It’s a (target word)”. The PowerPoint presentation was timed. Each picture was presented twice (42 pictures x 2 times = 84 pictures). The participants produced all the words correctly. None of the words were replaced or skipped. At the end of the recording session the participants were

asked what phenomenon was tested in the experiment. None of the participants guessed that word-final obstruents were the focus of the investigation. All participants filled in a language background questionnaire at the end of the recording session.

3.4. Analysis

The acoustic analysis was performed on the target words using the Praat software (Boersma & Weenink, 2011). Four durational measurements were collected from each token: (1) preceding vowel; (2) closure; (3) voicing into closure; (4) release. Then, voicing ratios were calculated for each word-final consonant using the following formula.

$$\text{Voicing ratio} = \frac{\text{Duration of voicing into closure}}{\text{Duration of closure}} \times 100$$

Voicing ratios showed how much voicing each consonant retained. Theoretically, a fully voiced consonant would have a voicing ratio of 100%, whereas a fully voiceless consonant a ratio of 0%.

A statistical analysis was performed on the measurements obtained from Russian learners and American English native speakers to identify the differences between the two groups and to establish which acoustic cues the Russian learners used to distinguish between voiced and voiceless consonants in English.

4. Results

The pilot study was a 2 X 2 factorial design. The independent variables were the feature [voice] (voiced, voiceless) and L1 (Russian, English). The dependent variables were the durational measurements, including (1) preceding vowel duration; (2) closure duration; (3) voicing into closure duration; (4) release duration, and voicing ratios.

Voicing ratios were calculated for each participant, and each group as a whole using the durations of closure and voicing into closure measurements. A two-way

ANOVA was run on the voicing ratios with the between-participants factors of L1 and voice. The significance cut off point was set at $p < .05$. The main effects of voice, $F(1,284) = 1067.02, p < .001$, and L1, $F(1,284) = 14.39, p < .001$, were significant. The Russian and American English speakers had different voicing ratios in their productions of voiced and voiceless English consonants. There was also a significant interaction between voice and L1, $F(1,284) = 29.16, p < .001$. The Bonferroni post hoc tests revealed that American English speakers ($M = 92, SD = 17$) had significantly ($p < .001$) higher voicing ratios than Russian speakers ($M = 70, SD = 35$) for voiced consonants. There was no significant difference found for voiceless consonants ($p = .257$).

Table 3

Mean voicing ratios (%) and standard deviations (in parentheses) in word-final English consonants for Russian and American English speakers

Russian speakers	Voiced	Voiceless	Am. English speakers	Voiced	Voiceless
R1	62 (15)	3 (6)	A1	99 (2)	2 (4)
R2	88 (27)	11 (9)	A2	81 (26)	1 (3)
R3	99 (2)	6 (14)	A3	100 (0)	1 (4)
R4	30 (34)	0 (0)	A4	86 (18)	0 (0)
Total:	70 (35)	5 (10)	Total:	92 (17)	1 (3)

Although as a group the Russian learners produced voiced consonants significantly differently from American English speakers, Table 3 shows that individual performance of some learners was similar to that of the American native speakers of English. For example, Russian learner R2 had a voicing ratio of 88%, which surpassed the performance of two American native speakers A2 (81%) and A4 (84%). Another Russian learner R3 had almost an ideal voicing ratio of 99%. However, a Russian learner R4 had a voicing ratio of only 30%, which means that he substantially devoiced voiced

stops word-finally. Although the learner has spent 8 years in the United States, he reported using Russian 70% of the time. It seems that his first language phonology has a strong effect on his English pronunciation.

In order to identify how Russian and American English speakers encoded voicing and whether the differences between Russian and American English speakers were significant, a series of two-way ANOVAs were run on the four measurements with the between-participants factors of voice and L1 (Table 4).

Table 4

Mean durational measurements (ms) and standard deviations (in parentheses) for Russian and American English speakers' productions of English consonants word-finally

Durations	Voiced consonants			Voiceless consonants		
	Russian	English	Δ	Russian	English	Δ
Vowel	147 (53)	170 (42)	23**	145 (39)	130 (26)	15*
Closure	74 (13)	87 (26)	13**	101 (26)	103 (36)	2
Voicing	52 (29)	80 (29)	28***	4 (9)	1 (3)	3
Release	110 (72)	60 (29)	50***	118 (66)	50 (31)	68***

Note. Δ reflects the average difference for each measure between productions by

American and Russian speakers. * $p < .05$, ** $p < .01$, *** $p < .001$.

The main effect of voice was found significant on vowel durations, $F(1, 284) = 18.76, p < .001$, closure durations, $F(1, 284) = 47.61, p < .001$, and voicing into closure durations, $F(1, 284) = 639.21, p < .001$. The main effect of L1 was found significant on voicing into closure durations, $F(1, 284) = 23.778, p < .001$, and release durations, $F(1, 284) = 89, p < .001$. There were significant interactions between voice and L1 on vowel durations, $F(1, 284) = 15.62, p < .001$, voicing into closure durations, $F(1, 284) = 38.52, p < .001$ and a marginally significant interaction on closure durations, $F(1, 284) = 3.78, p = .053$. The Bonferroni post hoc tests revealed that Russian learners and American

English speakers differed in vowel durations for voiced consonants ($p = .001$) and for voiceless consonants ($p = .029$); in closure durations ($p = .003$) and voicing into closure durations ($p < .001$) for voiced consonants only. Although there was no significant interaction between voice and L1 on release durations, the main effect of L1 on release durations shows that Russian and American English speakers differed significantly in their production of release irrespective of the feature [voice].

The Bonferroni post hoc tests also revealed what acoustic cues Russian and American English speakers used to encode the feature [voice]. Russian learners of English produced significantly ($p < .001$) shorter closure durations before voiced consonants than before voiceless and significantly longer ($p < .001$) voicing into closure durations before voiced consonants than before voiceless consonants (Figure 1).

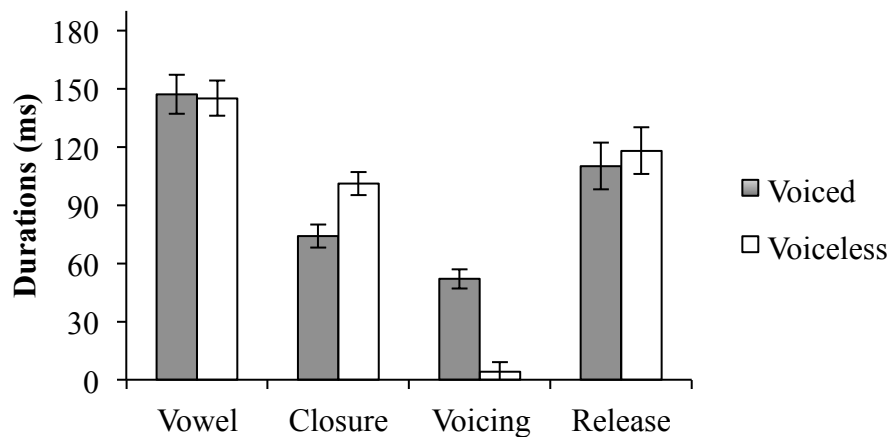


Figure 1. Mean durational measurements (ms) for Russian speakers' productions of English consonants word-finally. Error bars show the 95% CI.

American English speakers also manipulated closure durations and voicing into closure durations to differentiate between voiced and voiceless consonants: closure durations were significantly ($p = .001$) shorter before voiced consonants and voicing into

closure was significantly ($p < .001$) longer before voiced consonants. However, American English speakers also produced significantly ($p < .001$) longer vowel durations before voiced consonants than before voiceless consonants (Figure 2). Neither Russian learners of English nor American English native speakers used release durations as an acoustic cue to encode the feature [voice]. These findings along with the other results presented in this section will now be discussed in light of the research questions guiding the current investigation.

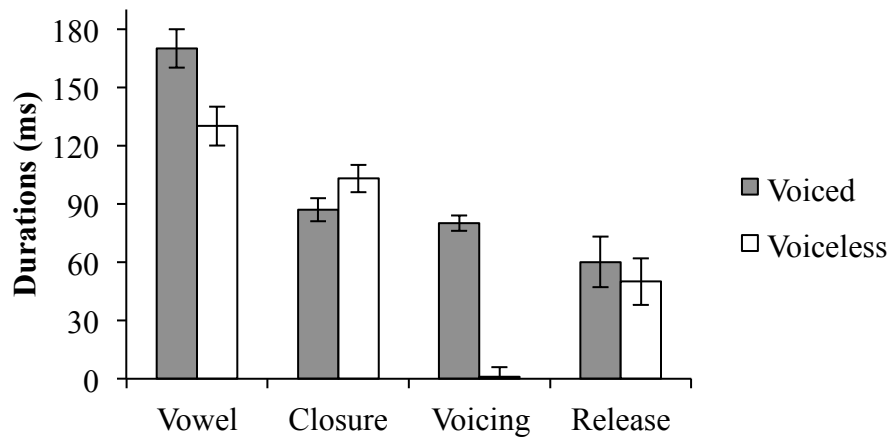


Figure 2. Mean durational measurements (ms) for American English speakers' productions of English consonants word-finally. Error bars show the 95% CI.

5. Discussion

The results of the pilot study suggest that it is possible for proficient late learners of English who are native speakers of Russian to acquire the feature [+voice] word-finally in English. Although the Russian participants as a group produced less voicing than the American native speakers, fifty percent of the Russian participants behaved in the range of American native speakers. Those Russian participants who reported less amount of L1 use in their daily lives in the US conformed more closely to American

native speaker rates than those who used Russian extensively. Further research is needed to confirm whether the amount of L1 use is indeed a confounding variable in the acquisition of word-final voicing. Thus, the first hypothesis that proficient Russian learners of English who have been immersed in the second language environment can produce voicing contrasts word-finally similarly to native speakers of American English was partially supported.

The second hypothesis that Russian learners of English use vowel duration as the main acoustic cue to distinguish voiced and voiceless consonants word-finally in English was not supported. The Russian learners only manipulated the durations of closure and voicing into closure to distinguish voiced and voiceless consonants word-finally in English, whereas the American English speakers manipulated vowel durations, closure durations and voicing into closure (Table 5).

Table 5

Summary of acoustic cues used by Russian and American English speakers in their production of English consonants word-finally

Durations	Russian speakers	American English speakers
Vowel		✓
Closure	✓	✓
Voicing into closure	✓	✓
Release		

As mentioned previously, vowel duration is a very important acoustic cue to differentiate voiced and voiceless consonants in English. A study by Raphael (1972) claims that American English listeners perceive English consonants as voiceless when they are preceded by vowels of shorter duration and as voiced when they are preceded by vowels of longer duration. Voicing into closure duration also has some effect on how

American English listeners categorize voiceless and voiced consonants, “although it is minor compared to that of vowel duration” (Raphael, 1972, p. 1301). The results of the current study suggest that Russian learners of English mainly rely on voicing into closure duration and closure duration to encode the feature [voice]. As a group, Russian learners of English achieved a voicing ratio of 70% (i.e., voicing into closure duration divided by closure duration and multiplied by 100) for word-final voiced stops. Consequently, the question arises as to how American English listeners categorize consonants produced by Russian speakers with a voicing ratio of 70% without differences in the duration of the preceding vowels. Future research is needed to determine at what voicing ratio consonants are perceived as voiced. One such experiment could use the productions of Russian learners as perception stimuli while a group of American English listeners serve as judges to categorize the stimuli as voiced or voiceless consonants and rate their goodness of fit. This analysis would help determine how much voicing ratio a consonant should possess for American English listeners to categorize it as voiced.

The pilot study also shed light on some flaws in the materials used to elicit data that have to be eliminated before the proposed study. First of all, ambiguous pictures should be replaced. Although they were tested on American native speakers prior to the pilot study, there were a few that consistently caused problems. They should also be tested on Russian native speakers to determine whether they elicit familiar English words. Secondly, another cycle should be added so that participants have to produce each word three times. Russian participants devoiced more during the second cycle than the first. Even the Russian learners who demonstrated good results had devoiced consonants in the second cycle. The participants likely found it boring and monotonous to produce

84 similar sentences. As a result, their control of English pronunciation might have started to decrease somewhere in the middle of the task. Another explanation might be that some Russian learners thought that the target phenomenon was related to vowel production and even if they did monitor their production, their main focus was on the wrong target. Thus, in a future study with a third cycle, the productions of the first cycle can be discarded as they may be the most monitored.

A number of improvements are necessary for the selection of target words. This study used words with both tense and lax vowels. Six words out of 18 had a tense vowel: 'seed', 'beet', 'cheek', 'sheep', 'loop' and 'cube'. Two words out of these six had tense vowels before voiced consonants and four words had tense vowels before voiceless consonants. Since in English tense and lax vowels differ not only in quality but also in quantity, it might have had an effect on the vowel durations in general. Russian has only five vowels and no quantity distinction. Therefore, some Russian learners of English often merge a tense-lax category of /i - ɪ / or /u - ʊ/ into the respective Russian categories of /i/ and /u/. In order to avoid any possible confounding factor of vowel duration due to the tense-lax distinction, only lax vowels should be included. The cost of this change will be the necessity of adding another mid vowel /ʌ/. Otherwise, a list of real words that would suit other requirements will not be possible. Another improvement to consider is controlling for the initial consonant to facilitate the process of measuring. Having liquids /l, ɹ/ or a glide /w/ as the initial consonants creates a lot of confusion because co-articulatory effects of preceding liquids and glides make identification of the vowel onset difficult. Given that problem, only words with the initial obstruents will be chosen

whenever possible. Allowing a vowel /ʌ/ will bring more lexical variation to choose from.

6. Conclusion

The ability to distinguish voiced and voiceless consonants in production is not freely given to native speakers of neutralizing languages, such as Russian. This pilot study suggests that while there are proficient Russian learners of English who can acquire the voicing contrasts word-finally, there are others who are severely affected by their native phonology. Despite a prolonged stay in the second language environment, learners who do not use the target language extensively are less likely to acquire the voicing contrast word-finally. Moreover, even those learners who produce word-final voiced consonants use acoustic cues differently from native speakers. The proposed study on the acquisition of word-final voicing by native speakers of neutralizing languages has the potential to answer questions about the acquisition of marked phonological processes and whether grammars that nonnative speakers develop can be nativelike. It can also provide insight into the effect of extralinguistic factors, such as the length of residence in the second language environment or amount of first language use, on the acquisition of phonological processes.

References

- Boersma, P. & Weenink, D. (2011): Praat: Doing phonetics by computer (Version 5.2.43) [Computer software]. Retrieved from <http://www.fon.hum.uva.nl/praat/>
- Bongaerts, T., Summeren, C., Planken, B. & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447-465.

- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22, 129-159.
- Davies, M. (2008). *The corpus of contemporary American English: 450 million words, 1990-present*. Retrieved from <http://corpus.byu.edu/coca/>
- Dinnsen, D. & Charles-Luce, J. (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics*, 12, 49-60.
- Dmitrieva, O., Jongman, A. & Sereno, J. (2010). Phonological neutralization by native and non-native speakers: The case of Russian final devoicing. *Journal of Phonetics*, 38, 483-492.
- Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, 27, 315-330.
- Flege, J. & Fletcher, K. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91, 370-389.
- Flege, J., Frieda, E. & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, 25, 169-186.
- Flege, J., Munro, M. & MacKay, I. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125-3134.
- Fourakis, M. & Iverson, G. (1984). On the 'incomplete neutralization' of German final obstruents. *Phonetica*, 41, 140-149.
- Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Oxford: Blackwell.
- Jassem, W. & Richter, L. (1989). Neutralization of voicing in Polish obstruents. *Journal of Phonetics*, 17, 317-325.

- 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). Retrieved from <https://ruor.uottawa.ca/en/handle/10393/22809>.
- Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.
- Mack, M. (1982). Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *Journal of the Acoustical Society of America*, 71, 173-178.
- Patkowski, M. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics*, 11, 73-89.
- Piroth, H. & Janker, P. (2004). Speaker-dependent differences in voicing and devoicing of German obstruents. *Journal of Phonetics*, 32, 81-09.
- Port, R. & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17, 257-282.
- Port, R. & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455-471.
- Pye, S. (1986). Word final devoicing of obstruents in Russian. *Cambridge Papers in Phonetics and Experimental Linguistics*, 5, 1-10.
- Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *The Journal of the Acoustical Society of America*, 51, 1296-1303.

- Reetz, H. & Jongman, A. (2009). *Phonetics: Transcription, production, acoustics and perception*. Oxford: Wiley-Blackwell.
- Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House/Harper & Row.
- Slowiaczek, L. & Dinnsen, D. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, 13, 325-341.
- Shrager, M. (2012). Neutralization of word-final voicing in Russian. *Journal of Slavic Linguistics*, 20, 71-99.
- Simon, E. (2010). Phonological transfer of voicing and devoicing rules: Evidence from L1 Dutch and L2 English conversational speech. *Language Sciences*, 32, 63-86.
- Smith, B., Hayes-Harb, R., Bruss, M. & Harker, A. (2009). Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics*, 37, 257-275.
- Warner, N., Jongman, A., Sereno, J. & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32, 251-276.