

Maja Marković*

University of Novi Sad

Faculty of Philosophy

Novi Sad, Serbia

PHONETIC FEATURE INTERPRETATION IN SECOND LANGUAGE PHONOLOGY: THE ACQUISITION OF THE FINAL OBSTRUENT VOICING CONTRAST

Abstract

Although languages may share certain phonological features, their phonetic realizations can be variable. The classical example is the difference in VOT as the underlying factor of the binary distinction in voicing. In this paper, we examine the voicing contrast of English final obstruents (plosives and fricatives) in the production of native speakers of Serbian. The paper reports the results of a production experiment, in which 14 relatively proficient speakers of English as L2 read eight monosyllabic words (minimal pairs) inside frame sentences. Although both English and Serbian contrast final voiced and voiceless obstruents, the results indicate that L2 speakers in our study have not fully mastered the phonetic realization of word-final voicing in English, but greatly retain the phonetic realization of voicing typical of their L1.

Key words: voicing, obstruents, pre-fortis clipping

* E-mail address: majamarkovic@ff.uns.ac.rs

1. Introduction

“Foreign accent” is the term that subsumes various deviations in the pronunciation of learners of a foreign language (L2). As has been well documented in numerous research studies, it often stems from the differences between the phonology of the learners’ native language (L1) and that of the target language. The tendency of L2 learners to use the phonetic and phonological features and patterns from their L1 while acquiring L2 is often referred to as ‘transfer’, the term originally proposed by Selinker (1969). Transfer is easily recognized with L2 sounds or contrasts which are non-existent in L1, such as the sounds /θ/ and /ð/, or the contrasts /e/ and /æ/ in English, or /u/, /y/ and /i/ in French. For the learners whose native language does not have these sounds or contrasts, acquiring English or French phonology would involve learning the new sounds and acquiring the new contrasts, or forming new categories, alongside the more familiar ones. The failure to do so often leads to substitutions or misinterpretations by learners.

A different situation occurs if the same phonological contrast exists in two languages. Although it may appear as the same, a contrast may be realized in alternative ways, but the learners will often assume that it functions in the same way in L1 and L2. Such is the difference in the realization of voiced and voiceless pairs of obstruents in English and Serbian, which is dealt with in this study. Although both English and Serbian have the distinction of voiced and voiceless stops, fricatives and affricates in all positions in the word, their phonetic realization involves different articulatory strategies, and hence different acoustic and auditory features in the two languages. This research focuses on the strategies of producing English word-final obstruents in minimal pairs, where one word ends in a voiced, and the other in a voiceless obstruent. The classes of obstruents investigated in this study are stops and fricatives.

The results of the research are interpreted within the framework of Flege’s Speech Learning Model (SLM; Flege 1995), in particular, in relation to hypothesis 6, which states that a phonetic category established for an L2 sound may be based on different phonetic features, or feature weights, from the features used by native speakers. The present research also looks into the other hypotheses and postulates outlined in the same study.

2. Voicing feature in English and Serbian obstruents

Voicing feature generally involves the presence or absence of vocal fold vibration during the articulation of a sound, referred to as the system of laryngeal features. Most obstruent sounds in the two languages studied in this paper maintain a binary contrast, namely [+/- voice].¹ However, the system of laryngeal features in English and Serbian is considerably different. The most broadly studied class of sounds in this respect are stop sounds, or plosives, in the word initial position. While English voiceless stops are typically aspirated in the stressed position, they are unaspirated in Serbian. On the other hand, word initial Serbian voiced stops are produced with pre-voicing – vocal folds vibrate during the closure phase of stop production, while in English initial voiced stops, vocal fold vibration starts only after the release phase. Jansen (2004) terms languages such as English ‘aspirating languages’, while Serbian belongs to the group of ‘voicing languages’. This phonetic distinction in the realization of the laryngeal features is related to the contrast in Voice Onset Time (VOT), established by Lisker and Abramson (1964). In terms of VOT, the contrast of word initial stops in Serbian is that between pre-voiced stops, with negative VOT values (voiced), and short-lag stops, where voicing begins shortly following the release (voiceless) (Čubrović 2011, Jakovljević 2012, Jerotijević Tišma 2017). English, on the other hand, has the contrast between short-lag (voiced) and long-lag (voiceless) stops. The lag in VOT in English voiceless stops is accompanied by a breathy sound following the release phase, known as aspiration.

An interesting aspect of voicing in all obstruents is that while voicelessness is relatively stable, voicing is rather variable, depending on the distribution of obstruents in the word. Under specific circumstances, voiced obstruents undergo the process of devoicing, i.e. they are phonetically voiceless. English obstruents are said to be partially voiced in word initial position, due to the relatively delayed VOT, fully voiced in the intervocalic position, but are completely devoiced in the word-final or pre-fortis position (Cruttenden 2014). The results of the study on a number of acoustic correlates of final obstruent devoicing in American English reported by Dmitrieva (2014) show that devoicing is not the same in all obstruents. While lenis stops and affricates are accompanied by a short

¹ With the exception of the fricative /h/ in English, which has no voiced counterpart, and the Serbian fricative /x/ and affricate /ts/, which also lack the voiced equivalent.

period of voicing in the word-final position, lenis fricatives are produced as completely voiceless.

While devoicing in English has been extensively studied, no systematic research has been done on devoicing in Serbian. According to Petrović and Gudurić (2010) and Bakran (1994), phonologically voiced obstruents retain some degree of phonetic voicing word-finally in Serbian and Croatian. However, the mechanism of devoicing has not been systematically studied in Serbian, and we therefore believe that this study may shed some light on this phenomenon.

3. Final obstruents

3.1. Pre-fortis clipping

Word-final position, which is the focus of this research, is particularly vulnerable to a number of consonant features. Stop sounds in this position can be optionally unreleased, while the contrast in phonetic voicing is suspended in all final obstruents. However, phonological contrast between voiced and voiceless consonants is still retained both in English and Serbian. Since English final voiced obstruents are voiced only phonologically, a different contrast and another set of terms may be more appropriate, i.e. the contrast between fortis (strong, phonologically voiceless) and lenis (weak, phonologically voiced) obstruents.

The question that inevitably arises here is how the contrast between fortis and lenis word-final consonants is achieved, since most of the phonetic contrast is erased in this position. As has long been known, fortis consonants have the property of shortening, or clipping the vowel preceding them, and the information on the vocalic duration acts as a much more reliable acoustic cue to the listener on the nature of the final consonant (House and Fairbanks 1953, Denes 1955, House 1961). The reduction of vowel duration before voiceless consonants is a universal tendency of all languages (Chen 1970), but is more marked in some languages than others. English exhibits a very strong tendency towards this kind of reduction, often referred to as pre-fortis clipping (term introduced by Wells 1990). According to Wiik (1965), long vowels can be twice longer in open syllables or in front of voiced (lenis) obstruents or sonorants than when followed by a voiceless consonant at the same speaking rate. From

the point of view of perception, this difference acts as the major auditory cue facilitating the discrimination between the final consonants (Collins and Mees 2005: 51; Cruttenden 2014: 206; Roach 1998: 33; Yavas 2006: 58). Obviously, English elevates the universal phonetic property of vowel clipping to a higher level in the phonological system, so that it has been ‘phonologized’, to use the term proposed by Kohler (1984), in order to convey the distinction in voicing.

3.2. Other characteristics of voicing in final obstruents

Leaving the vowel duration aside, one might (wrongly) assume that the English final voiced obstruents and their voiceless pairs do not differ whatsoever. However, this is not the case, since fortis consonants tend to be more prominent than the lenis ones under the same speaking rate and at comparable intensity. This prominence is primarily reflected in the temporal features of final obstruents. Final fortis fricatives and affricates are characterized by a significantly longer duration than their lenis equivalents (Denes 1955; Cole and Cooper 1975), while final fortis stops are characterized by a longer occlusion phase than their lenis equivalents (Lisker 1957; Port 1979). Thus the duration of the fortis/lenis consonant is inversely related to the duration of the preceding vowel, and it was found that the ratio between the duration of the vowel and the duration of the consonant is the main auditory signal facilitating the recognition of the contrast for native speakers of English (Derr and Massaro 1980; Flege and Hillenbrand 1986).

3.3. Final obstruents in Serbian

Although Serbian has not been studied as extensively as English, it is known that final obstruents are not completely devoiced (Petrović and Gudurić 2010: 104–105, Bakran 1996: 84). This means that the friction in lenis fricatives or the occlusion phase of final lenis stops may be accompanied by vocal fold vibration either for its whole duration, or for a certain percentage of its duration. The study by Marković and Jakovljević (2016), which deals with the acquisition of pre-fortis clipping in English by native speakers of English, shows that Serbian also exhibits duration reduction before voiceless obstruents. The reduction is more marked with long vowels than with short ones, but much less marked than in English.

To our knowledge, no study has been conducted on the duration of the final obstruents in Serbian, or the ratio between the vowel and consonant duration in relation to consonant voicing and its significance for the perception of consonants.

In addition to vowel duration and the presence of voicing in the final fricative, it is also known that the final voiceless fricative in Serbian tends to be pronounced with higher intensity, and that the intensity peak is concentrated around the central portion of the fricative (Jerotijević Tišma 2017 and references therein).

A typical distinction between the final /s/ and /z/ in Serbian is seen in Figure 1, which shows the pronunciation of the words *bos* and *knez* by one of the subjects of the present study. The vowels of both words are phonologically long.²

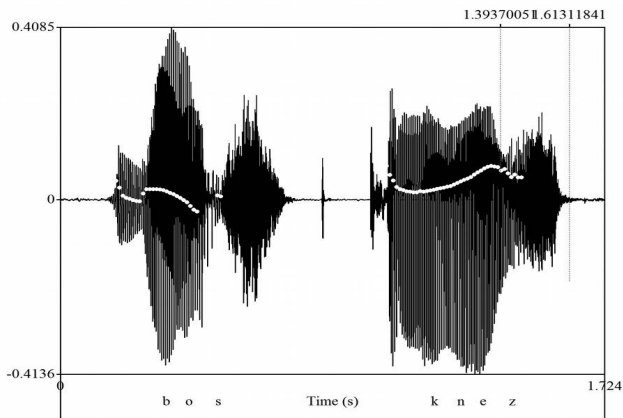


Figure 1. The waveform of the Serbian words *bos* and *knez*

The duration of the vowel in the word *bos* in Fig. 1 is 205ms, and of that in *knez* is 254ms, the duration of the voiceless fricative /s/ is 242ms, and of the voiced /z/ it is 214ms. Obviously, Serbian also exhibits the tendency towards reducing the duration of the vowel before voiceless fricative, and towards longer duration of voiceless fricatives. The white dots in Figure 1 indicate the presence of voicing, which continues into the voiced /z/

² Serbian vowel duration is traditionally related to pitch accents, which can be long falling, long rising, short falling or short rising. Monosyllabic words can only have falling pitch accent, so the words analyzed here both have long falling accents.

for approximately 30% of the duration of friction. Another feature that distinguishes the voiceless from the voiced fricative in Serbian is related to their energy level and distribution. In terms of maximum intensity level, it can be seen that the voiceless fricative has somewhat higher maximum, but the more striking difference between them is that the peak of energy is concentrated around the mid portion of the voiceless /s/, but evenly distributed during the voiced /z/.

Figure 2 shows the difference between the voiceless and voiced pair of stops /t/ and /d/ in the words *pet* and *bled* (both with long falling accent), also pronounced by one of the subjects of this study.

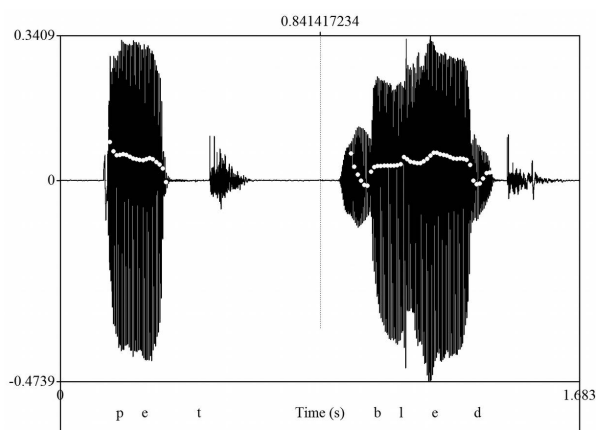


Figure 2. The waveform of the Serbian words *pet* and *bled*

The duration of the vowel preceding the voiceless stop /t/ is 184ms, while the same vowel lasts for 219ms before the voiced stop /d/. The duration of occlusion in /t/ is 147ms, and in /d/ it lasts for 116ms. The white dots in Fig. 2 show the presence of vocal fold vibration, which goes well into the occlusion phase and lasts for 81ms (70% of occlusion duration).

To sum up, Serbian exhibits some phonetic similarities and some differences in the realization of final fricatives and stops when compared to English. The features they share are the duration of the vowel preceding the voiceless/voiced consonant, the duration of the voiced/voiceless fricative and occlusion in stops, and consequently the ratio in the duration of the vowel and voiceless/voiced obstruent. Yet, all of these features are much more marked in English than in Serbian and are therefore perceptually the most salient for the distinction between final fortis and lenis consonants.

The two languages differ in the presence/absence of voicing in the final voiced obstruent. While Serbian retains a period of voicing in the final voiced obstruent, English has no phonetic voicing in the final lenis obstruent. It is probable that the English durational characteristics have been exaggerated in order to compensate for the lack of the phonetic distinction in voicing.

Acknowledging that the temporal relations of the vowel and final consonant act as the major signal for distinguishing between the voicing contrast in English raises the question of the acquisition of this property by L2 learners whose native language does not share the same mechanism.

4. Previous research in final obstruent acquisition

In this section we shall give a brief overview of studies on the acquisition of final obstruents relevant for our research. Several studies have dealt with the acquisition of the final obstruent acquisition by learners whose L1 neutralises voicing contrast in the word-final position. Among those, McAllister (2007) investigated the realization of the English /s/-/z/ contrast in a production study of speakers of Swedish. This study was based on the production of 17 speakers of Swedish, relatively proficient in English. The results indicate that the voiced fricative is rather poorly acquired, and that even the most successful speakers did not employ the same articulatory/acoustic mechanisms as those typical of native speakers' pronunciation. Another study, which investigated the perception and production of the word-final voicing contrast by German learners of English, was conducted by Smith et al. (2009). It also showed that the German speakers produced fewer and less robust acoustic cues to word-final voicing than English speakers. Several studies were conducted on the acquisition of English final obstruent contrasts by speakers of Czech, another language without final voicing contrast. The results reported in Fejlová (2013) and in Skarnitzl & Šturm (2016) show that the Czech speakers did not sufficiently exploit duration to cue the identity of the word-final obstruent. The study by Skarnitzl and Šturm (2014) showed that the Czech subjects of their research tended to produce both the fortis and lenis obstruents without any phonetic voicing. In the study on the acquisition of the final /t/-/d/ contrast by speakers of Chinese, English and Spanish, it was found that native Chinese speakers of English, most of whom were inexperienced in English, could more accurately distinguish between voiced and voiceless

stop consonants in word-initial than in word-final position (Flege et al. 1992).

The studies dealing with the acquisition of English consonants by learners whose first language is Serbian are of relatively recent date. A comprehensive study by Jerotijević Tišma (2017) deals with the general acquisition of all English consonants by native speakers of Serbian in all positions in the word and in different speaking styles/contexts. It included several perception and production tasks and tested 72 students of English at University of Kragujevac. The results, including those related to the word-final position, point to a high level of transfer from the learners' L1, manifested as substitutions and assimilation of target sounds to the categories of L1. The author also found that the learners were generally successful in differentiating phonological differences, but not the subtle phonetic ones. The studies by Čubrović (2011, 2012) and Jakovljević (2012, 2014) dealt with the distinction in VOT in the acquisition of fortis and lenis plosives by Serbian speakers. Marković and Jakovljević (2016) report on the results of a study of pre-fortis clipping of English diphthongs in the production of Serbian learners. The results point to partial, but generally insufficient level of attainment of this temporal feature before fortis obstruents.

5. Present study

5.1. Hypotheses

The goal of the present study was to investigate how native speakers of Serbian acquire several phonetic parameters that underlie the voicing contrast in English, in particular, vowel duration, the duration of the final fricative/occlusion of stop consonants and the duration of voicing in the final obstruent.

Based on the previous research on some of the above parameters, we hypothesized that native speakers of Serbian at relatively high levels of proficiency in English may not generally have acquired at least some of the phonetic exponents of the voicing contrast in L2, but resort to the more familiar cues of their L1. The hypothesis of our research is of interest in relation to Flege's Speech Learning Model (SLM, Flege 1995) in several respects. Firstly, our study tests the 'feature' hypothesis (hypothesis 6),

which states that an L2 phonetic category formation may be blocked if listeners do not discern a mismatch in the phonetic features used to signal a contrast in L1 and L2, i.e. if the learner's representation is based on different phonetic features, or feature weights (Flege 1995: 239). Relevant to our study are also hypotheses 2 and 3, which state that a new phonetic category can be established for an L2 sound if it differs phonetically from the closest L1 sound. If the perceived phonetic similarity is not significant, an L2 sound is more likely to become identified with an equivalent L1 sound, which is the implication of hypothesis 3. In line with these, we also hypothesize that the similarity between the L1 and L2 contrast in phonological voicing will hinder the acquisition of the L2 contrast, and that it will be more or less identified with the contrast typical of the learners' L1.

The goal of the study was to test whether and to what extent advanced learners of English have mastered the above-mentioned phonetic exponents of the phonological distinction of voicing in their interlanguage and discuss some implications for foreign language phonological acquisition.

5.2. Participants and procedure

The participants of this study were 14 first year students who were recorded in the soundproof studio at the Faculty of Philosophy, University of Novi Sad. The students participated in the study on a voluntary basis, and none of them reported any hearing impairment or speech defect. All of them had learned English for at least eight years prior to the study. Materials used for this research were monosyllabic minimal pairs containing a diphthong in the nucleus and an obstruent in the coda.³ The words selected for this study were *mace*, *maize*, *pace*, *pays*, *height*, *hide*, *tight* and *tide*. Target minimal pairs were placed within frame sentences, and both words occurred in the second and third stressed positions (W1 and W2), but were followed by another stressed word (*Did you say W1, W2 or W3?* or *I said W1, W2 and W3.*). Since the target words were presented as parts of lists, all participants read them with a short break before the next word. This ensured that

³ Although the participants' native language, Serbian, does not contain diphthongs, we opted for diphthongs for two reasons: firstly, since the durational differences in the pre-fortis and pre-lenis environments are generally more marked in long vowels, especially diphthongs, and secondly, because our previous research showed the highest degree of acquisition of this durational difference in diphthongs among our participants, especially the ones analyzed here /eɪ/ and /aɪ/ (Marković 2012; Marković and Jakovljević 2016).

potential voicing due to the medial (intervocalic) position should not occur. The second and third stressed positions were taken for the even amount of stress, and the final position was not taken into consideration because of the marked final lengthening of both the vowel and consonant, which would render invalid results for this study. The same sentences were read by a male native speaker of Standard British English, whose recordings were used as a control.

The recordings were analysed using Praat software for speech analysis (Boersma and Weenink 2011). The analysis included the measurements of vowel duration, the duration of friction for the fricatives /s/ and /z/, the duration of occlusion for the stops /t/ and /d/, and the duration of glottal pulses (voicing) during friction/occlusion. In the segmentation of fricatives, we followed the procedures recommended by Kent and Read (1992) and Peterson and Lehiste (1960). We also checked the distribution of energy in the fricative pair, but did not include the numerical data of this parameter in this study. The obtained data were statistically analysed for the mean values of duration, standard deviations and ratio between the duration of vowels and consonants. T-test was conducted to determine the statistical significance of difference between the relevant sets of data. The data were also qualitatively compared to samples of minimal pairs in Serbian, produced by the same participants.

6. Results

The overall results of duration measurements are shown in Table 1. The data therein show mean duration (in milliseconds) of the vowel preceding the final fortis/lenis fricative or stop, the duration of the fricatives, the duration of the occlusion phase in stops and the duration of voicing in the final obstruent during friction or occlusion. It also contains the information on standard deviations from the mean duration values in the subjects' production. The grey shaded fields show the duration values in the material read by the control native speaker.

	[-V] fricative		[+V] fricative		[-V] stop		[+V] stop				
	V	C	V	C	voicing	V	occlusion	V	occlusion	voicing	
Dur (ms)											
L	255.3	196.3	290.8	185.2	52.9	252.7	106.5	306.1	89.6	53.9	
SD	41.3	44.4	43.1	34.8	24.1	34.6	30.8	31.5	32.8	30.4	
NS	193	271	360	175	0	204	75	374	34	11	

Table 1. Duration of vowels (V), final consonants (C), occlusion and voicing in fortis ([-V]) and lenis ([+V]) environments. L = learners, NS = native speaker

6.1. Duration of vowels

While the control native speaker consistently pronounces the vowel before all of the fortis consonants with remarkably shorter duration than before the lenis ones, the learners are far less consistent.⁴ The ratio of the pre-lenis to pre-fortis mean vowel duration of the native speaker is 1.85, while that of the learners is only 1.18. The results of vowel duration measurements are shown in the graph in Figure 3.

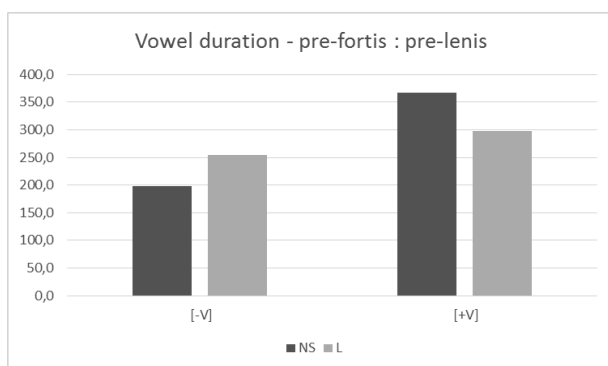


Figure 3. Vowel duration before lenis and fortis obstruents

The results of the t-test conducted show that the learners tend to produce longer vowels before lenis than fortis consonants ($p < 0.01$), although the variability among their duration values and ratios between pre-fortis and pre-lenis duration is remarkable, as can be seen from the data on standard deviations (Table 1). The standard deviation in the ratio between pre-lenis and pre-fortis duration of the learners' vowels is 0.17. Notably, only one of the learners (subject 10) has the duration ratio of pre-fortis and pre-lenis vowels comparable to that of the control speaker in front of the final fricative pair (1.73), and one before the final stop pair (subject 9, ratio 1.61). Four learners produced shorter vowels in the pre-lenis environment, three of whom did so before the final lenis fricative, and one before the lenis stop. Upon the exclusion of the values obtained by the speakers who failed to produce adequate durations, the ratios of the other speakers of

⁴ The values obtained for the native speaker were expected, since they are in line with the findings well documented in the literature and in previous research. We therefore assume that the data obtained in our study are relevant, although they rely on the pronunciation of a single native speaker.

pre-lenis and pre-fortis vowel duration are somewhat higher than in the learners' L1, i.e. Serbian (c.f. Sovilj-Nikić 2007; Jakovljević 2012, 2014; Marković and Jakovljević 2016).

Comparing the duration of vowels before fricatives and before stops, the overall results indicate somewhat lesser variability in the duration values before fortis and lenis stops and their ratios than before fricatives. However, t-test showed no statistically significant difference of the vocalic duration in the two environments ($p > 0.05$).

6.2. Duration of final fricatives

The duration of English final [-V]/[+V] fricatives is inversely related to the duration of the preceding vowel. In other words, word-final fortis fricatives have significantly longer duration than the lenis ones. The duration of the final fortis fricative /s/ pronounced by our control speaker is 271ms, while that of the lenis fricative /z/ is 175ms, the ratio between the fortis and lenis being 1.55. The mean duration of /s/ in the production of our subjects is 196ms, whereas that of /z/ is 185ms, with the mean ratio 1.06 between them. With the exclusion of two subjects who pronounced the fortis and the lenis fricative with the duration ratio comparable to that of the control native speaker (subjects 3 and 10), no other subjects exhibited the tendency typical of native speakers of English. T-test showed no significance between the duration of the final fortis and lenis fricatives in the production of our subjects ($p > 0.05$). This finding was unexpected, since the subjects' L1 also has the tendency of longer voiceless fricatives. The reason may be the fact that the difference in consonant duration is not so marked in the learners' L1 as in L2, and compared to vowel duration, it is not as audible.

McAllister (2007 and references therein) states that the ratio between the duration of the vowel and that of the consonant can be taken as the main signal by which English native speakers recognize the contrast between word-final voiced and voiceless fricative. Our study showed that the ratio between the duration of the vowel and the fortis fricative /s/ produced by the control speaker is 0.7 (193ms : 271ms), while that of the vowel and the following lenis fricative /z/ is 2.1 (360ms : 175ms). The mean ratio of the duration of the vowel and the fortis fricative in our subjects' production is 1.3 (SD=0.4, 255.3ms : 196.ms), and that of the vowel and lenis consonant is 1.6 (SD=0.3, 290ms : 185.2ms). T-test

again showed no significance in the two duration ratios in our subject's production ($p > 0.05$).

6.3. Duration of occlusion

With word-final stop sounds, the control native speaker had significantly longer occlusion phase in the production of fortis stops than the lenis ones (mean duration 75ms : 34ms). The data on the mean duration of the subjects' occlusion phase (Table 1) show a similar tendency, although the difference between fortis and lenis stop occlusion duration is not as prominent as in the native speaker's production (106.5ms : 89.6ms). Although several learners (subjects 6, 11 and 12) have a similar ratio between the fortis and lenis stop occlusion to that of the control speaker, the other subjects have highly inconsistent results. The comparison of the two sets (fortis and lenis occlusion duration) shows statistical significance at the level $p < 0.05$, but not at $p < 0.01$ ($p = 0.01834$). Another notable difference between the native speaker's pronunciation and that of the learners is that the occlusion phase is generally longer in the learners' production in all stop consonants. This may be due to the overall slower speech tempo of foreign speakers, although it was not the case with the duration of the final fricatives analyzed. Similar results were found by Tišma-Jerotijević (2017).

6.4. Duration of voicing

English phonologically voiced obstruents are generally characterized by a tendency of devoicing in the word-final position. The data of our study show a complete lack of glottal pulses in the control speaker's production of the final fricative /z/ (voicing duration = 0, Table 1). In contrast, all the subjects of our study pronounced /z/ accompanied by a period of voicing, which on average lasts for 52.9ms (29% of the duration of the fricative). The duration of voicing is also highly variable among the subjects, ranging from 14% to 66% of fricative duration.

The devoicing process is not as marked with the final stop sound /d/ in the pronunciation of the control native speaker. Our data show that the control speaker's occlusion phase is accompanied by a short period of voicing (11ms, or 33% of occlusion). In the production of our subjects, word-final /d/ has a significantly longer duration of voicing in the occlusion

phase than produced by the native speaker. The mean duration of voicing in the subjects' pronunciation of /d/ was 53.9ms, or 63% of the duration of occlusion. Four learners (subjects 1, 3, 4 and 9) even pronounced the lenis stop with voicing throughout the occlusion phase. One of the subjects pronounced /d/ with a very brief period of voicing (7%), but failed to produce the vowel before /t/ and /d/ differently, and as a result, the words *tight* and *tide* were completely indistinguishable.

7. Discussion and concluding remarks

English is a language in which voicing contrast of obstruents is not neutralized in the final position, but has a strong tendency of devoicing of lenis obstruents. Serbian similarly retains the contrast in voicing in the final position, and also has lesser marked voicing features in the phonologically voiced final obstruents than in other positions, complying with the universal tendency of all languages. However, 'voicing' is a general term which subsumes different phonetic features, and does not straightforwardly refer to the activity of the vocal folds. Final position is peculiar in that the voicing contrast has at least three phonetic features that combine into the single phonological feature referred to as 'voicing'. The ones we looked into in this study are the temporal features, involving the duration of the preceding vowel, the duration of the final consonant, and the duration of glottal pulses (voicing) in the obstruent. In addition to these, the difference may be expressed in terms of intensity, either as the intensity level or intensity distribution. Although intensity was not analyzed in this study, we noticed that it may also have contributed to the distinction between voiced and voiceless fricative pair /s/-/z/ in our subjects' production. Figure 4 shows the pairs *maize-mace* produced by subject 2 (a) and *pays-pace* produced by the control speaker (b).

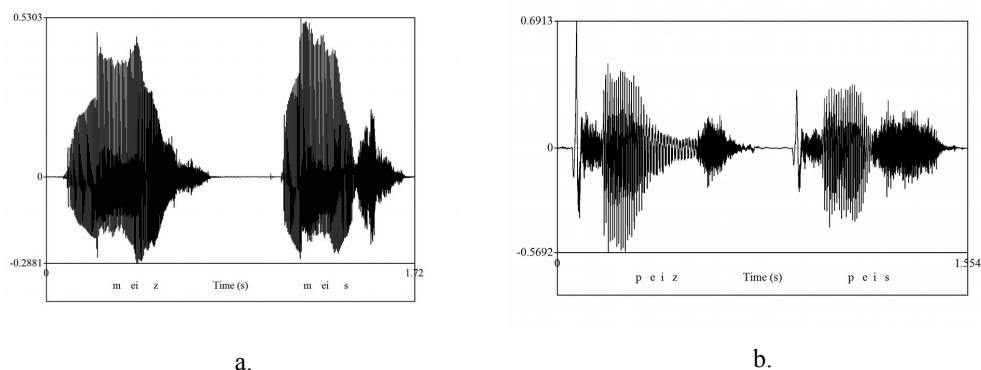


Figure 4. The waveform of the words *mace* and *maize* pronounced by a learner (a) and *pays* and *pace* pronounced by the control native speaker

The image shows that the main difference between the fortis and lenis fricative sound in the learner's production is in the intensity level, where the voiceless fricative has a markedly higher amplitude, and the concentration of energy is around the mid portion of friction.⁵ In the pair of words produced by the native speaker, the energy of the voiceless fricative is not significantly higher, and is distributed evenly for the whole duration of friction. The main distinction in the native speaker's production is achieved by the difference in duration, the fortis /s/ being much longer than the lenis /z/. Since the intensity features noticed in our study were hugely variable and even speaker specific, we decided to exclude these data from the analysis and return to those in a future, more extensive study of intensity. Yet, intensity cannot be completely disregarded as another phonetic distinction of voicing in the two languages analyzed.

Returning to the temporal features as a dominant phonetic exponent of the voicing contrast, our results strongly indicate that they are not fully acquired by our subjects, but are overall transferred from the learners' L1. Although we did not conduct a systematic analysis of the same phonetic features in Serbian, it is known from previous studies that Serbian also exhibits duration reduction of vowels preceding voiceless consonants (Sovilj-Nikić 2007, Marković 2012, Marković and Jakovljević 2016).

⁵ The same tendency of energy distribution of Serbian speakers was noticed in the research of Tišma Jerotijević, who attributes the difference between distributed energy in English fricatives and concentrated energy peaks in the Serbian ones to the difference in the place of articulation in the two languages.

Despite the variability that our subjects exhibited, none of them had acquired all the phonetic features typical of the native speaker's pronunciation. As was noted earlier, subject 10 had the ratio between the duration of the vowel before the fortis and lenis fricative approximately the same as the native speaker. Since vowel duration is probably the most significant auditory cue of the voicing contrast, this subject's pronunciation sounded the most native-like. However, the same subject had the longest period of voicing in the lenis fricative /z/, which indicates that he acquired only some of the phonetic features relevant for the distinction. Similar results were found by McAllister (2007), whose best sounding subjects also had a significant voicing period in the final lenis fricative, and their success to produce a native-like sounding contrast was due to the rearranged temporal patterns in the vowels alone.

Vowels preceding fortis and lenis final stops exhibit the same durational patterns as before fricatives. We noticed, however, that the ratio between the fortis and lenis occlusion phase in the production of our control speaker was not as consistently high as in the fricative pair. In addition, it was observed that the native speaker also had a period of voicing in the occlusion phase of the final lenis stop. Although literature on English states simply that final stops are devoiced, it appears that the devoicing is not as complete as in fricatives. Our study corroborated the similar findings of the study of final obstruent devoicing in American English reported by Dmitrieva (2014), who found that final lenis fricatives are completely devoiced, unlike final lenis stops. Still, the voicing period during the occlusion phase in the production of our control speaker is significantly shorter than that of the subjects, who again resort to voicing as the main distinguisher between /t/ and /d/. The results show that even four subjects retain voicing throughout the occlusion phase of /d/, while the others on average have a much higher percentage of voicing duration than was found in the native speaker's pronunciation.

To return to our initial hypotheses, we can conclude that they have been mainly confirmed in this study. First, the results show that the learners have mainly retained the phonetic features of the voicing contrast typical of their L1, and that the existence of the voicing contrast in Serbian blocked the acquisition of the phonetic features typical of L2, resulting in the identification of the contrast with that of the learners' L1, which is in line with Flege's hypothesis 6. The most prominent phonetic feature of the voicing contrast, i.e. the duration of the preceding vowel, is by far

the best acquired one, as far as our results indicate. In other words, the subjects had more marked difference in the duration of the pre-lenis and pre-fortis vowel than is typical in their L1. This corroborates Flege's (1995) hypotheses 2 and 3, which states that the greater the perceived phonetic dissimilarity in an L1 and L2 contrast or category, the more likely it is to be earlier acquired by learners. Statistical analysis showed that the other phonetic characteristics, such as the duration of the final consonant (or occlusion) and the presence of voicing in the final lenis obstruent is mainly transferred from the learners' L1.

Finally, since some of the phonetic features in our subjects' production approximate those of native speakers', we can conclude that the acquisition of a complex contrast, involving multiple phonetic features, is not an all-or-nothing scenario, but probably shows the evolving character of L2 phonological acquisition.

This small-scale study revealed the necessity of further exploration not only in the field of second language phonological acquisition, but also of more extensive studies of voicing features both in Serbian and in English. Even though Serbian has been understudied in the issues of voicing at this moment, it appears that different English dialects should also be researched in more detail to come up with more precise descriptions of final voicing distinctions in this language.

References

- Bakran, J. (1996). *Zvučna slika hrvatskoga govora*. Zagreb: Ibis grafika.
- Boersma, P. and D. Weenink (2011). *Praat: doing phonetics by computer* (Version 5.2.28). (28 June 2011) <<http://www.praat.org>>.
- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22, 129–159.
- Cole R. A. and W. E. Cooper (1975). Perception of voicing in English affricates and fricatives. *Journal of the Acoustical Society of America*, 59, 1280–1287.
- Collins, B. and I. M. Mees (2005). *Practical Phonetics and Phonology*. London: Routledge.
- Cruttenden, A. (2014). *Gimson's Pronunciation of English*. London: Edward Arnold.

- Čubrović, B. (2011). Voice Onset Time in Serbian and Serbian English. *ELOPE: English Language Overseas Perspectives and Enquiries*, 8(1), 9–18.
- Čubrović, B. (2012). The Evolution of a Phonetic Phenomenon: The Case of Voice Onset Time in Serbian Intermediate EFL Learners. In: T. Paunović & B. Čubrović (eds.), *Exploring English Phonetics*, Newcastle upon Tyne: Cambridge Scholars Publishing, 47–55.
- Denes, P. (1955). Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America*, 27, 761–764.
- Derr, M. A. and D. W. Massaro (1980). The contribution of vowel duration, f₀ contour, and frication duration as cues to the /juz/-/jus/ distinction. *Perception & Psychophysics*, 27, 51–59.
- Dmitrieva, O. (2014). Final voicing and devoicing in American English. *Journal of the Acoustical Society of America*, 136(4), 2174–2174.
- Fejlová, D. (2013). Pre-fortis shortening in fluent read speech: A comparison of Czech and native speakers of English. *AUC Philologica 1/2014, Phonetica Pragensia XIII*, 101–112.
- Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In: W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language speech research*. Timonium MD: York Press, 233–277.
- Flege, J. E. and J. Hillenbrand (1986). Differential use of temporal cues to the [s-z] contrast by native and non-native speakers of English. *Journal of the Acoustical Society of America*, 79, 508–517.
- Flege, J. E., M. J. Munro and L. Skelton (1992). Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America*, 92, 128–43.
- House, A. S. (1961). On vowel duration. *Journal of the Acoustical Society of America*, 33, 1174–1178.
- House, A. S. and G. Fairbanks (1953). The influence of consonantal environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America*, 25, 105–113.
- Jakovljević, B. (2012). VOT Transfer in Production of English Stops by Serbian Native Speakers. In: T. Paunović, T. & B. Čubrović (eds.), *Exploring English Phonetics*, Newcastle upon Tyne: Cambridge Scholars Publishing, 31–46.
- Jakovljević, B. (2014). The strength of articulation of the members of the voice contrast in English and Serbian. In: S. Guduruć (ed.), *Jezici i kulture u vremenu i prostoru 3*. Novi Sad: Filozofski fakultet, 481–489.

- Jansen, W. (2004). *Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology Approach to English, Hungarian, and Dutch*. PhD dissertation. University of Groningen.
- Jerotijević Tišma, D. (2017). *Percepcija i produkcija konsonantskog sistema u englesko-srpskoj međujezičkoj fonologiji*. Doktorska disertacija. FILUM, Univerzitet u Kragujevcu.
- Kent, R. D. and C. Read (1992). *The Acoustic Analysis of Speech*. San Diego: Singular Publishing.
- Kohler, K. J. (1984). Phonetic Explanation in Phonology: The Feature Fortis/Lenis. *Phonetica*, 41(3), 150–74.
- Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language*, 33, 42–49.
- Lisker, L. and A. S. Abramson (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 348–422.
- Marković, M. (2012). *Usporedna proučavanja vokala engleskog i srpskog jezika: između univerzalnog i specifičnog*. Novi Sad: Filozofski fakultet.
- Marković, M. and B. Jakovljević (2016). Acquiring vocalic quantity and quality in L2: The Acquisition of vowel clipping in English by Advanced Serbian Learners. *Zbornik Matice srpske za filologiju i lingvistiku*, LIX/2, 97–108.
- McAllister, R. (2007). Strategies for realization of L2-categories. English /s/-/z/. In O. S. Bohn & M. J. Munro (eds.), *Language experience in second language speech learning. In honor of James Emil Flege*. Amsterdam: John Benjamins, 153–166.
- Peterson, G. E. and I. Lehiste (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America*, 32(6), 693–703.
- Petrović, D. i S. Gudurić (2010). *Fonologija srpskoga jezika*. Beograd: Institut za srpski jezik SANU.
- Port, R. F. (1979). The influence of tempo on stop closure duration as a cue for voicing and place. *Journal of Phonetics*, 7, 45–56.
- Roach, P. (2008). *English Phonetics and Phonology: A Practical Course*. Cambridge: CUP.
- Selinker, L. (1969). Language transfer. *General Linguistics*, 9, 67–92.
- Skarnitzl, R. and P. Šturm (2014). Assimilation of voicing in Czech speakers of English: The effect of the degree of accentedness. *Research in Language* 12(2), 199–208.
- Skarnitzl, R. and P. Šturm (2016). Pre-fortis shortening in Czech English: A production and reaction-time study. *Research in Language*, 14, 1–14.

- Smith, B. L., R. Hayes-Harb, M. Bruss and A. Harker (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.
- Sovilj-Nikić, S. (2007). *Trajanje vokala kao jedan od prozodijskih elemenata u sintezi govora na srpskom jeziku*. Magistarska teza. Novi Sad: Fakultet tehničkih nauka, Univerzitet u Novom Sadu.
- Wells, J. C. (1990). Syllabification and Allophony. (20 October 2015) <<http://www.phon.ucl.ac.uk/home/wells/syllabif.htm>>.
- Wiik, K. (1965). *Finnish and English vowels*. Turku: Turku University.
- Yavas, M. (2006). *Applied English Phonology*. Oxford: Blackwell.

ACKNOWLEDGEMENT

The research presented in the paper is supported by the technological project “The Development of Dialogue Systems for Serbian and Other South Slavic Languages” (TR32035) funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia, and the project CABUNS, funded by the Secretariat of Higher Education and Scientific Research.

Received: 17 June 2019

Accepted for publication: 18 July 2020

Maја Марковић

ИНТЕРПРЕТАЦИЈА ФОНЕТСКИХ ОБЕЛЕЖЈА У ФОНОЛОГИЈИ СТРАНОГ ЈЕЗИКА: УСВАЈАЊЕ КОНТРАСТА ОБЕЗВУЧЕНИХ ФИНАЛНИХ ОПСТРУЕНАТА

Сажетак

Иако различити језици могу имати иста релевантна фонолошка обележја, њихове фонетске реализације се могу међусобно разликовати. Типичан пример јесте разлика у времену наступа звучности (ВОТ), фактору који одређује бинарну опозицију између звучних и безвучних консонаната. У овом раду бавимо се

опозицијом звучних и безвучних опструената (пловива и фрикатива) у финалној позицији у изговору енглеских речи код изворних говорника српског језика. У раду се саопштавају резултати експеримента продукције, у ком је 14 говорника који уче енглески језик као страни изговарало осам једносложних речи (минималних парова) у реченичном оквиру. Иако и енглески и српски језик имају контраст између звучних и безвучних опструената у финалној позицији у речи, резултати показују да испитаници нису у потпуности овладали фонетским карактеристикама овог контраста у енглеском, већ у великој мери примењују карактеристике типичне за њихов матерњи језик.

Кључне речи: звучност, опструенти, скраћивање пред безвучним опструентима