ISSN 1451-5342
eISSN 1820-5682

Naučno-stručni časopis za jezik, književnost i kulturu
broj 16 • 2018. • godina XVI • Beograd

PHILOLOGIA

Academic journal for language, literature and culture
Volume 16 • 2018 • year XVI • Belgrade

Specijalni tematski broj: fonetika i fonologija
Special thematic issue: phonetics and phonology

Izdavač • Published by
PHILOLOGIA

Kontakt adresa uredništva:
Knez Mihailova 40, 11000 Beograd, Srbija
www.philologia.org.rs
Email: philologia.journal@gmail.com

SADRŽAJ/TABLE OF CONTENTS

Izdavački savet • Advisory Council

UVODNA REČ

Pred vama je novi broj časopisa *Philologia,* po prvi put koncipiran kao specijalni tematski broj u kome su izloženi vredni rezultati samostalnih istraživanja iz oblasti fonetike i fonologije. U šesnaestom po redu tomu časopisa *Philologia* svoje mesto je našlo sedam inovativnih fonetskih i fonoloških studija o fonetskim fenomenima u različitim jezicima sveta i iz perspektive govornika sa različitih jezičkih područja, što ovoj publikaciji daje poseban značaj. Odabrane fonetske studije koje su se našle u ovoj svesci časopisa usmeno su izložene na Četvrtom međunarodnom sastanku fonetičara engleskog jezika koji je od 30–31. marta 2018. održanom na Filološkom fakultetu u Beogradu (Fourth Belgrade International Meeting of English Phoneticians – BIMEP).

U prvom radu, profesorka Elizabet Ziga iz perspektive artikulatorne fonologije i tradicionalnih fonoloških teorija obeležja rasvetljava procese vezanog govora na primerima iz engleskog kao stranog jezika sa stanovišta izvornih govornika korejskog jezika. Kako bi se pružio što bolji uvid u procese vezanog govora, odabrani primeri iz engleskog (kao maternjeg) i ruskog, takođe su uključeni u analizu. Ova inovativna studija je posebno važna iz razloga što analizom govornih navika neizvornih govornika objašnjava univerzalne fonološke procese kao što su nazalizacija i gubljenje zvučnosti.

U narednom članku Andrej Bjelaković pruža rezultate akustičke analize engleskih vokala u produkciji izvornih govornika srpskog jezika. Akustička analiza spektralnih karakteristika vokala pokazuje da ispitanici načelno ne zamenjuju vokale engleskog jezika vokalima srpskog, ali su vokali stranog jezika samo kompromisne replike vokala koje ne dosežu formantske vrednosti koje su karakteristične za izvorne govornike engleskog jezika.

Fonetska studija Dimitre Dimitriu bavi se načinom artikulacije rotičkog glasa engleskog jezika kod govornika kiparskog grčkog. Autorka proučava realizacije glasa /r/ u različitim pozicijama u slogu i u produkciji različitih grupa govornika kiparskog grčkog i zaključuje da nijedna grupa ispitanika nije u potpunosti usvojila ovaj glas, kao i da su efekti sistema prvog jezika vidni u njegovoj realizaciji u engleskom.

Marlisa Homel u svom radu govori o ulozi ortografije i njenom uticaju na percepciju kod holandskih srednjoškolaca koji uče engleski kao strani jezik. Testovi percepcije pokazuju da pravopisne karakteristike engleskog jezika otežavaju percepciju segmenata engleskog jezika, kao i da učenici „nove" glasove teže prepoznaju od poznatih.

U svom radu Višnja Josipović Smojver razmatra potencijalnu ulogu fonoloških teorija u podučavanju engleske fonetike na univerzitetskom nivou. Autorka na odabranim primerima ilustruje moguću primenu fonoloških teorijskih okvira kao što su grkljanska fonologija, teorija otvora, artikulaciona teorija i teorija optimalnosti u podizanju fonološke svesti studenata o problemima u izgovoru sa kojima se susreću u engleskom jeziku.

Komparativnom analizom fonoloških sistema engleskog i španskog jezika Marija Alisija Maldonado nastoji da rasvetli problematična mesta u izgovoru sa kojima se najčešće susreću izvorni govornici španskog koji uče engleski jezik. Varijetet španskog koji se opisuje u ovom radu govori se u delu Argentine koji se graniči sa Urugvajem. Autorka daje inovativne i korisne predloge u cilju poboljšanja izgovora engleskih segmentalnih i suprasegmentalnih pojava.

Nikoleta Stojkova zatvara ovaj tom svojim radom o upotrebi slabih oblika engleskih funkcionalnih reči u produkciji izvornih govornika bugarskog jezika na univerzitetskom nivou. Upotreba slabih oblika u engleskom jeziku ispituje se na kraćem vezanom tekstu i daju se preporuke za uspešno savlađivanje njihovog izgovora.

Zahvalnost, najpre, upućujemo svim autorima. Bez njihovog velikog angažovanja ne bi bilo no ovog toma časopisa. Recenzentima, koji su svojim stručnim kritičkim pristupom podigli kvalitet štampanih radova, posebno smo zahvalni. Redakcija časopisa poziva sve potencijalne autore na saradnju. Prilozi se mogu slati na i-mejl adresu našeg časopisa, a sve pojedinosti o rokovima, tehničkim uputstvima i smernicama za pisanje radova mogu se naći na veb-stranici: www.philologia.org.rs.

Glavni i odgovorni urednik
Prof. dr Biljana Čubrović

A WORD FROM THE EDITORIAL

Philologia is a peer-reviewed academic journal whose primary objective is to promote and advance research in the humanities. The journal comes out annually in electronic edition.

Philologia publishes articles, critical essays, book reviews, interviews, conference reports grouped into the following sections: Linguistics, Applied Linguistics, Literary Studies, Cultural Studies, Translation Studies, Scientific Interviews, Conference Reports and Book Reviews. All previous issues are available at: www.philologia.org.rs

This year's thematic issue of *Philologia* brings seven original scientific articles providing innovative perspectives in the fields of phonetics and phonology. We hope you will find the papers stimulating and ispirational.

The Editorial Board is endlessly obliged to the Reviewing and Advisory Council which includes Serbian and international reviewers. The significance of their role can be better understood when taking into consideration the fact that only 65% of all papers sent to the Editorial Board has passed through the fine sieve of their careful reading and evaluation. Together with reviewers, the Editorial Board keeps implementing strict criteria when selecting papers for publication. This is our contribution to raising the quality of science and research.

We are also very much grateful to our colleagues who sent the papers to have them reviewed. Without their trust in the work of Editorial Board and Advisory Council, it would have been impossible to have this journal published.

Philologia Editorial Board

UDC: 811.111'342.9
811.531'342.9
DOI: https://doi.org/10.18485/philologia.2018.16.16.1

THE PHONETICS AND PHONOLOGY OF ENGLISH CASUAL SPEECH: LEARNING FROM L2 LEARNERS

ELIZABETH ZSIGA[1]
Department of Linguistics
Georgetown University
Washington, DC
USA

U radu se ispituju procesi „vezanog" ili „opuštenog" govora pri izgovoru drugog jezika, sa fokusom na govor korejskih učenika engleskog jezika. Polazi se sa stanovišta artikulatorne fonologije, koja tvrdi da su mnoge asimilacije i gubljenja u opuštenom govoru rezultat preklapanja artikulatornih gestova. Primerima iz engleskog i ruskog ilustrujemo preklapanje gestova. Ostali primeri potiču iz detaljnije fonetske studije procesa nazalizacije i jednačenja po zvučnosti u korejskom i engleskom sa korejskim akcentom (Zsiga 2011). Korejsko-engleski podaci pokazuju postepeno preklapanje gestova u jednačenju po zvučnosti, a u pojedinim slučajevima i prisustvo delimične nazalne asimilacije, što ide u prilog artikulatornoj fonologiji. Međutim, primećeni su i mnogi primeri kategoričke nazalne supstitucije. Tvrdimo da tradicionalniji fonološki pristup promene obeležja bolje objašnjava kategoričke promene. Da bi se u potpunosti objasnili svi rezultati, neophodne su i artikulatorna fonologija i tradicionalna fonologija zasnovana na fonološkim obeležjima.

Ključne reči: opušten govor, učenje drugog jezika, fonologija, artikulatorna fonologija, korejski, engleski s akcentom, nazalna asimilacija.

1. INTRODUCTION: ARTICULATORY PHONOLOGY AND CASUAL SPEECH

For many phoneticians who study English, interest in second-language speech is primarily practical, focused on how to change learners' pronunciation for the better. In contrast, this paper will take a more theoretical stance. Rather than focusing on what

1 Contact information: Elizabeth.Zsiga@georgetown.edu

phoneticians can teach second-language (L2) learners, this paper will focus on what phoneticians can *learn* from L2 learners. Because L2 pronunciation combines aspects of different linguistic systems, studying that interaction can provide insight into the structure of Language (with a capital L) in general. Just as physicists can learn about the basic structure of atoms by investigating what happens when particles collide, so the study of L2 speech production provides an opportunity to learn about the basic elements of Language, by examining native patterns that persist or new patterns that are created when language systems collide. This paper, specifically, will discuss acoustic data from Korean and Russian learners of English, with an emphasis on English "connected" or "casual" speech.

The term "casual speech" refers to a speech style that contains the optional assimilations and deletions that are more common in less formal pronunciations. These often occur at word boundaries, so this speech style can also be called "connected speech." Examples include final consonant deletion, as in "mashed potatoes" pronounced as "ma[ʃ-p]otatoes," or assimilations, as in "tin pans" pronounced as "ti[m-p]ans," or a combination, as in "sandwich" pronounced as "sa[m]wich." In understanding these kinds of alternative pronunciations, I follow the theory of Articulatory Phonology, expounded by Louis Goldstein and Catherine Browman in a series of articles beginning in 1986. The theory has been widely adopted and updated since then, including work summarized in this paper. (For overviews, see Browman/Goldstein 1990, 1992, Goldstein et al. 2006, Hall 2010, Zsiga 2011 and references therein.)

In the theory of Articulatory Phonology, the basic "atoms" of phonological organization are articulatory gestures: movements of articulators toward the goal of making a constriction in the vocal tract at a particular place with a particular degree of closure. Examples include labial closing, velum opening, or pharyngeal constriction. Lexical items can contrast in the presence or absence of gestures, gestural specification, and in the way the gestures are coordinated in time. Allophony is the result of differences in gestural organization.

Gestural organization can be graphically represented in a "gestural score," as shown in Figure 1. As in an orchestral score, the x-axis is time, and the different rows of the score indicate the parts played by different articulators, sometimes sequential, sometimes overlapping. In Figure 1A, the word "bad" is shown to have three gestures: labial closing, lowering and fronting of the tongue body, and closure of the tongue tip at the alveolar ridge. Note that the consonant gestures overlap in time with the gesture for the vowel. Because the labial closure and tongue body gesture begin at approximately the same time, the tongue body will fully reach the vowel "target" by the time the lips open. The vocal tract remains open for a time for the vowel, and then the tongue tip gesture for [d] begins as the vowel gesture ends.

In Figure 1B, a velum opening gesture has been added to the end of the word, creating "ban." (If the gesture had been added at the beginning, the word would be "mad.") The velum opening overlaps not only with the consonant, making a fully nasal [n], but begins during the vowel. This pattern of overlap creates partial vowel nasalization. The allophonic "rule" of vowel nasalization comes about because of this pattern of gestural timing.

In Figure 1C, a laryngeal opening gesture has been added, which turns "ban" into "pan." In the current model, the default state for speech is voicing, so only gestures for

*de*voicing (laryngeal opening) are indicated. Note that the laryngeal opening gesture extends into the beginning of the vowel. The result of this pattern of co-ordination is aspiration: a period of time after the labial closure has been released, during which the larynx remains open. Again, this allophonic "rule" comes about because of a specific pattern of gestural timing.

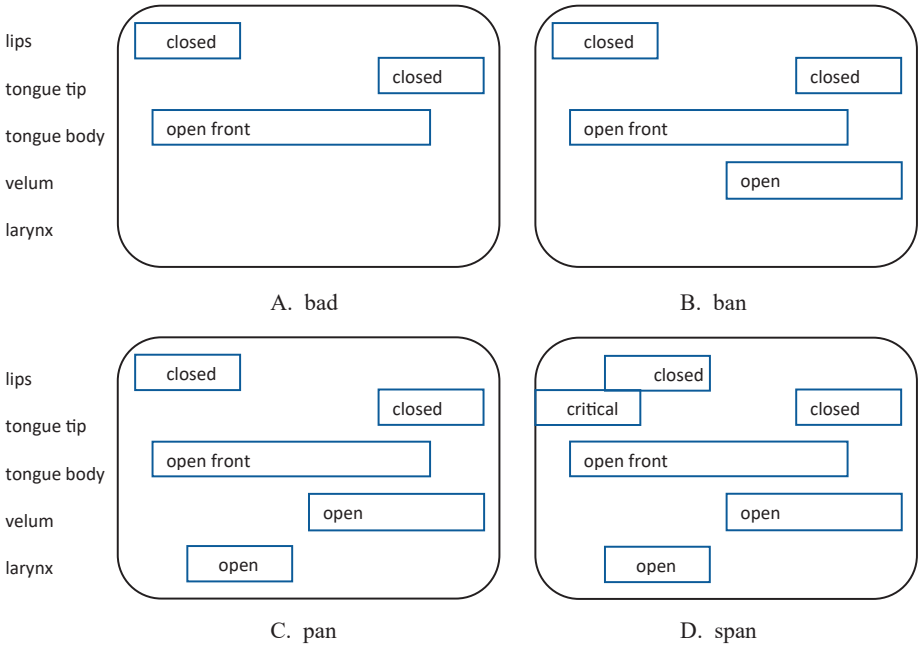
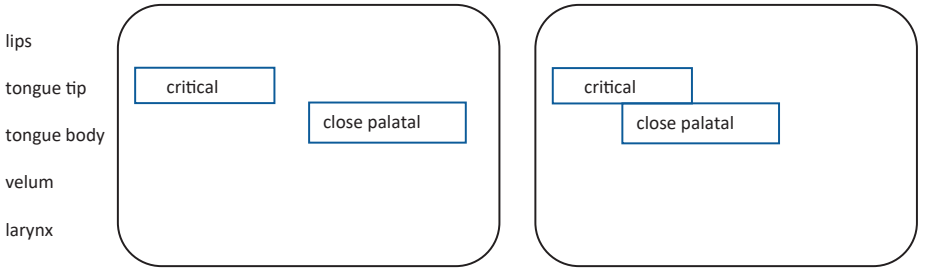


Figure 1. Gestural scores for the words "bad," "ban," "pan" and "span"

Finally, Figure 1D shows the word "span." A gesture for an alveolar fricative (a "critical" constriction of the tongue tip, meaning exactly the right constriction to create turbulence) has been added. In an initial consonant cluster, the two consonant gestures are re-organized: [p] moves somewhat to the right and [s] somewhat to the left, so that the laryngeal opening is centered in the middle of the cluster, not on any one gesture (Browman/Goldstein 1988). The result is that the laryngeal opening no longer extends into the vowel: [p] is unaspirated after [s].

Patterns of gestural overlap can also account for allophonic variation at word boundaries. For example, Zsiga (1995, 2000) argued that palatalization of [s] to [ʃ] in American English phrases like "miss you" is the result of gestural overlap. Figure 2A shows gestures for [s] (critical tongue tip) and [j] (palatal tongue body) as they might be organized in the words "miss" and "you" pronounced with a pause between. There is no overlap between the final [s] and the initial [j]. But in "connected speech" in English, the words overlap, as shown in Figure 2B. As a result, the two gestures blend together, producing a fricative that sounds most like [ʃ], though careful acoustic analysis shows that it is not exactly the same as [ʃ].



A. [s # j] with pause B. [s # j] with overlap

Figure 2. Gestural scores showing that palatalization in American English can be the result of gestural overlap

Figure 3 shows a spectrogram of the phrase "press you" pronounced in connected speech by a native English speaker (Zsiga 2000). The first part of the fricative is high-pitched, corresponding to [s]. The second part, however, shows overlap with [j]: note the high F2 and F3 formants visible on the spectrogram, co-extensive with the second half of fricative noise. The acoustic result of the overlap is lowered pitch, more similar to [ʃ]. The pitch lowering, however, is partial and gradual. Zsiga (2000) interprets the gradient change as evidence against a phonological rule substituting [ʃ] for [s], and in favor of an analysis based on gestural overlap.

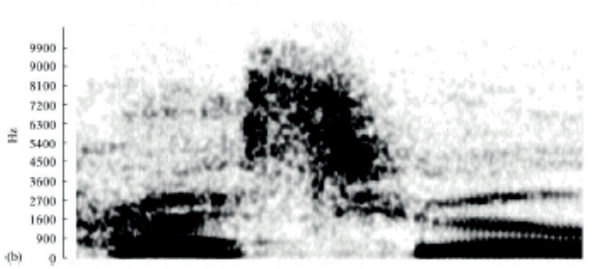


Figure 3. Spectrogram of the phrase "press you" pronounced in connected speech by an English speaker, showing gradient palatalization due to gestural overlap (Zsiga 2000: 85)

It is important to keep in mind, however, that patterns of gestural overlap are language specific. Figure 4 (also from Zsiga 2000) shows an [s#j] sequence in Russian (from the phrase /pas jejo/ "tended it.") For a Russian speaker, words do not overlap: the fricative /s/ is completed before the fricative /j/ begins. This pattern of overlap is not just a fact about /s/ and /j/: Russian speakers in general keep their words separate, and thus Russian does not show the type of connected speech assimilations typical of English. Note that palatalized [sʲ] is a different story: for this contrastive segment

the two gestures are simultaneous, and both are carefully maintained, a pattern of articulation that English speakers learning Russian find very hard to accomplish. English learners of Russian, however, are liable to pronounce the name "Boris Yeltsin" as Bori[ʃ] Yeltsin, transferring the English pattern of overlap to their Russian pronunciation.

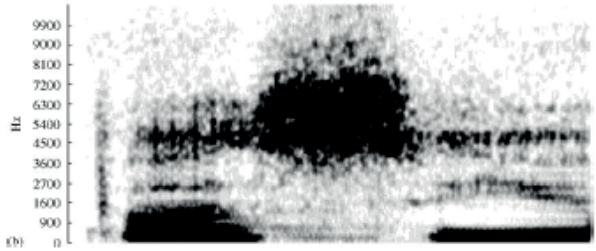


Figure 4. /pas jejo/, native Russian pronunciation (Zsiga 2000: 86)

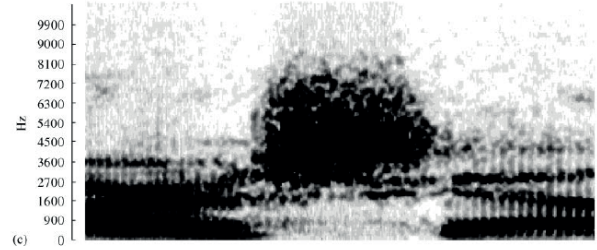


Figure 5. "Boris Yeltsin," pronunciation by a native English speaker (Zsiga 2000: 89)

Is the assimilation seen in American English palatalization, and other casual speech processes, phonology or phonetics? Traditionally, such alternations are described in terms of phonological rules. The lexicon specifies a basic form of each word, and then a set of rules indicate how to pronounce these words in different environments, by inserting, deleting, or changing one segment into another: [n] becomes [m] before a labial, or [s] becomes [ʃ] before a palatal, either within words or at word boundaries. Only then, after the phonological computation is completed, the speaker's brain sends instructions to the articulators to produce the specified sequence.

Articulatory Phonology argues that phonological rules such as these are not needed. According to the theory, the lexicon specifies different basic word forms that the speaker memorizes. Some of these stored lexical items may be semantically or syntactically related to each other, such as "press" and "pressure" or "balance" and "imbalance," but they are not derived from one another by phonological rule. For the variable and partial assimilations and deletions of connected speech, Articulatory Phonology explains them in terms of gestural reorganization and overlap as described

above, not segment substitution. The question then arises of whether there is space for "phonology" in Articulatory Phonology. Does a speaker make generalizations that are language-specific, but independent of lexical specification, "rules" that are more general than the pronunciation of specific words, but planned at a higher level than the unintended consequences of coarticulation?

To answer this question, we can learn by looking at the speech of language learners. L2 learners often transfer segmental pronunciation and allophony from the L1. What do they do with connected, casual speech? If the assimilations of casual speech are associated with specific lexical items, we would expect that no transfer from L1 to L2 would occur, because the lexical items in the two languages are different. If assimilations do transfer, and are found to be partial and gradient, that would be evidence for an account based in phonetic coarticulation and gestural overlap. The third possibility is that casual speech alternations do transfer, but look like categorical substitutions rather than gestural blending. If that is the case, then a phonological level, separate from both the lexicon and from articulatory organization – a level of productive phonological rules or constraints – must be posited.

2. CONNECTED SPEECH IN L2

Studies on connected speech in L2 are rather sparse, and results are not uniform. Most studies have found that learners often *don't* use the relevant rules of casual speech from either the L1 or L2. Instead they may use an "interlanguage" pattern that corresponds to neither. The usual interlanguage pattern seems to keep words separate, even if applying rules from the L1 would make L2 pronunciation more native-like.

For example, Weinberger (1994) reported on the speech of Mandarin speakers of English. In careful speech, Mandarin does not allow word-final obstruents. But in casual speech, Mandarin speakers can delete certain final vowels, leaving obstruents in word final position, so that /tòufu/ may be pronounced as [tòuf]. (Vowels are deleted in casual speech if they are stressless and toneless, and share the place of articulation of the preceding consonant: [u] after labials, [i] after dentals, and retroflex after retroflex.) Final epenthesis of vowels is found in neither L1 English nor L1 Mandarin. Nonetheless, Weinberger found that the same speakers who pronounce /tòufu/ as [tòuf] in their L1 pronounce "loaf" as [lofu] in L2 English. The casual speech rules that allow word-final [f] in the L1 don't transfer to the L2. Epenthesis takes place, even though word-final [f] would be allowed in both L1 and L2 if casual speech rules applied.

Cebrian (2000) reports a similar pattern for Catalan learners of English. He discusses two rules of Catalan: word-final devoicing, illustrated in (1), and regressive assimilation across word boundaries, illustrated in (2).

(1) Catalan word-final devoicing

/vaz/		/gos/	
va[z]os	"glasses"	go[s]os	"dogs"
va[s]	"glass"	go[s]	"dog"
/reb/		/escup/	
re[b]re	"to receive"	escu[p]ir	"to spit"
re[p]	"receives"	escu[p]	"spits"

(2) Catalan regressive voicing assimilation across word boundaries

/vaz/		/gos/	
va[z] gran	"big glass"	go[z] gran	"big dog"
va[s] petit	"small glass"	go[s] petit	"small dog"
/reb/		/escup/	
re[b] molts	"receives lots"	escu[b] molt	"spits a lot"
re[p] cartes	"receives letters"	escu[p] tot	"spits everything"

As (1) shows, voiced obstruents are devoiced in word-final position: /vaz/ "glass" is pronounced as [vas]. But in connected speech, as shown in (2), the word-final obstruent assimilates to the voicing of a following consonant: [vaz gran] "big glass," [vas petit] "small glass." Cebrian found that Catalan learners of English transfer word-final devoicing, but not cross-word voicing assimilation, to L2 English, as shown in (3).

(3) Catalan speakers' pronunciation of English phrases

wise guy → wi[s] guy
proud girl → prou[t] girl

In the cases of "wise guy" and "proud girl," transferring the Catalan rule of voicing assimilation would have resulted in the correct English pronunciation, but the Catalan speakers in Cebrian's study did not do that. Cebrian argues that the speakers were obeying an "interlanguage prosodic constraint" he terms "Word Integrity." The principle of word integrity requires that words in L2 be unconnected. Word Integrity "treats every word as a separate unit and prevents the synchronization of sounds belonging to different words" (2000: 19).

Zsiga (2003) also found evidence for word integrity in pronunciations by Russian learners of English. These speakers tended to release final consonants in clusters, for example pronouncing "make parts" as m[ek^{h} p^{h}]arts, rather than the more overlapped L1 English pattern that results in unreleased final consonants. But does word integrity always hold of L2 speech? A subsequent study, reported in Zsiga (2011) and summarized below, found that it does not.

3. AN EXPERIMENT ON NASALIZATION IN KOREAN ENGLISH

Zsiga (2011) recorded the speech of 12 native Korean learners of English. Six were classified as advanced L2 speakers, with a mean of 15.4 years of English instruction (usually beginning in middle school), and a mean of 4.1 years of residence in the United States. These were speakers with very high TOEFL scores or the equivalent, who were studying in degree programs at U.S. universities. The other six speakers were classified as intermediate. They also had many years of English instruction (mean of 8.5 years) but had lived in the U.S. for a year or less (mean of 5.9 months). They were for the most part family members of students or diplomats, and all were enrolled in intermediate-level ESL classes rather than degree programs. In addition to the 12 native Korean speakers, three native English speakers were recorded as controls.

The study focused on one aspect of Korean and Korean-English pronunciation: word-final voiceless stops followed by a word-initial nasal. In Korean, these stops undergo nasal assimilation, as shown in (4).

(4) Nasal assimilation in Korean

[pap] "rice"	[pam mekta] "eat rice"
[ot] "clothes"	[on man] "only clothes"
[jak] "medicine"	[jaŋ mekta] "take medicine"

The 12 speakers read sentences in Korean and English that contained stop#nasal sequences and nasal#nasal sequences, interspersed with other sentences that targeted obstruent#obstruent sequences, in randomized order. There were eight nasal-target sentences in Korean and 16 in English. Korean examples included [kimpa**p m**ekta] "eat sushi" and [o**t n**eta] "put clothes in," matched with structurally similar English phrases such as "kee**p M**att awake" and "bough**t n**ine." Korean and English were recorded in separate blocks, with the Korean sentences read first. Instructions were given in Korean by a native Korean-speaking research assistant, and the Korean sentences were written in Korean orthography. Three repetitions of each sentence were recorded. The three native English speakers read only the English sentences.

The recordings were then transcribed by two listeners, one native Korean, one native English, focusing on the word-final consonant. Acoustic measures were also taken, including consonant duration, duration of voicing, duration of nasalization, and presence or absence of audible release on the word-final consonant. The question to be answered was: how did the Korean speakers pronounce these word-final consonants? Did they follow the native Korean pattern, the native English pattern, or an interlanguage pattern that corresponds to neither L1 nor L2?

The native Korean tokens were analyzed to establish the L1 baseline. As predicted, in native Korean, 93% of the obstruent#nasal sequences were pronounced as nasal#nasal. An example is shown in Figure 6, which shows the syllables [on man] from underlying /os mantulta/ "make clothes." There was no measurable difference between underlying and derived [n#m]. A few tokens in the dataset (5%) were pronounced with the word-final consonant fully or partially voiced but not nasalized, and a few others

(2%) remained fully voiceless. The categorical change in 93% of tokens argues for a phonological substitution: In Korean, [n] is substituted for [t] and [m] for [p] when a nasal consonant begins the next word.

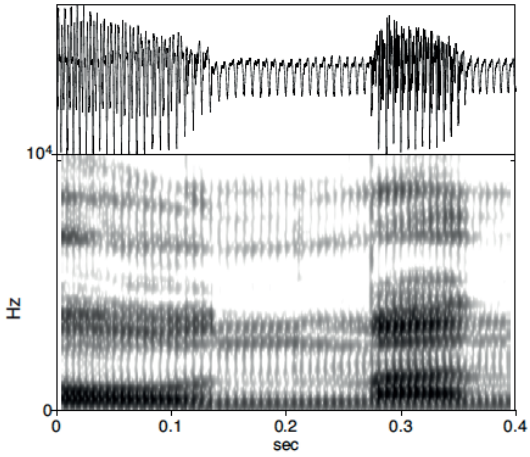


Figure 6. Spectrogram of [on man] from /os mantulta/, make clothes (Zsiga 2011: 307)

Examples of stop#nasal sequences in native American English are shown in Figure 7. Consistent with previous studies by Cohn (1993) and by Huffman (1989), 76% of the obstruent#nasal sequences produced by the native English speakers had voiceless closure with no release, usually with some glottalization on the preceding vowel. An additional 13% had creaky voicing throughout the closure. Only 11% showed an audible release between stop and nasal. None of the tokens produced by native English speakers had modal voicing or nasalization during the stop closure.

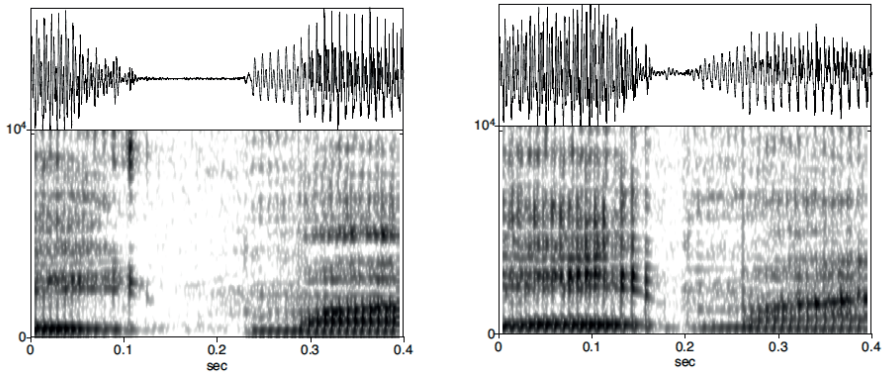


Figure 7. Spectrograms of two repetitions of the phrase "ate mine" by native English speakers (Zsiga 2011: 309)

In contrast to the consistency shown by native speakers of both languages, the realization of pre-nasal stops by Korean speakers of English was highly variable, as shown in Table 1.

percentage	realization	example spectrogram in:
15%	voiceless stop, released	Figure 8
49%	unreleased stop with gradient voicing	Figure 9
23%	full nasal	Figure 11
9%	partial nasal	Figure 12
4%	other (pause or mispronunciation)	

Table 1. Realization of word-final stops in pre-nasal position, Korean speakers of English

Though no speaker was completely consistent, each speaker did have a predominant or preferred pattern of pronunciation: four speakers produced a majority of tokens with nasal consonants, seven produced a majority of tokens with unreleased consonants, and one produced a majority of tokens with released final consonants. It is worth noting that there was no significant effect of level of instruction. The additional years of English instruction that the advanced speakers had received, and their high levels of proficiency in written English, did not make a significant difference in their pronunciation of these sequences.

Figure 8 shows an example spectrogram of a final voiceless stop with audible release, from the phrase "ate mine" by Speaker K8. This is the pattern that is predicted by the principle of Word Integrity proposed by Cebrian (2000) and given preliminary support in the Russian-English data collected by Zsiga (2003). These released final stops, although they are not typical of either L1 Korean or L1 American English, they do signal a clear separation between words. Tokens exhibiting word integrity were a minority in these data, however, at only 15% of tokens, mostly produced by only one speaker.

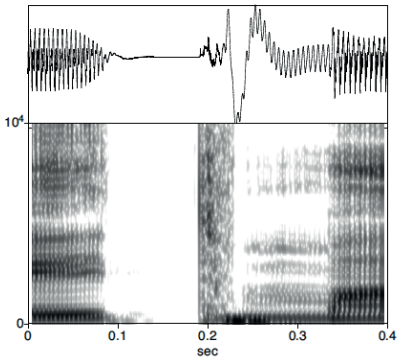
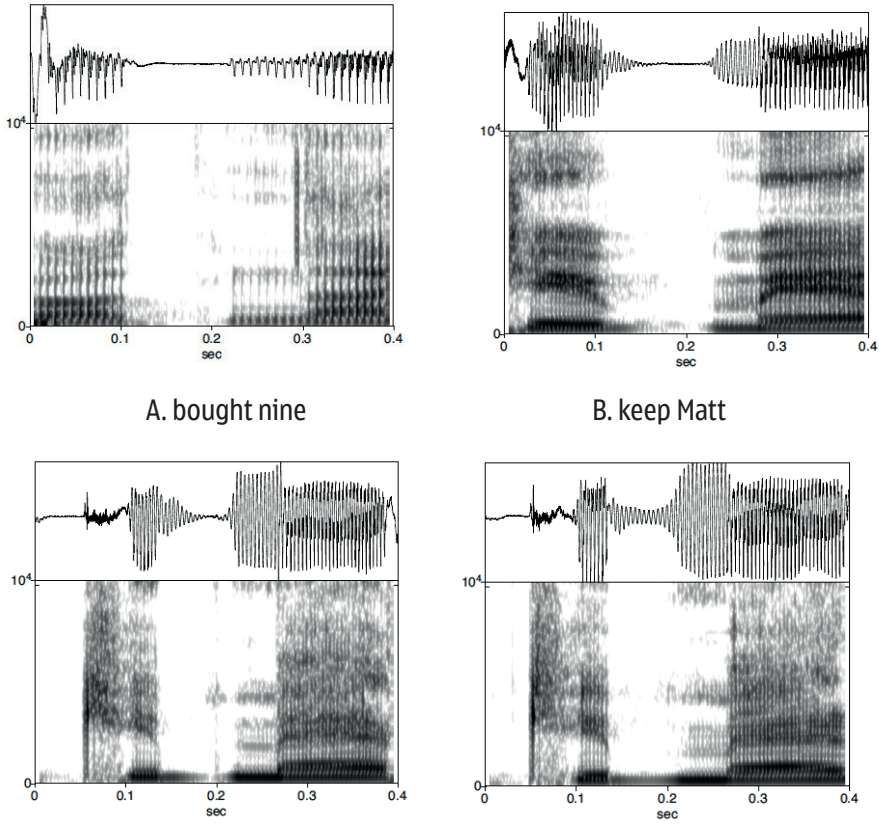


Figure 8. Spectrogram of the phrase "ate mine" in Korean English, with audibly-released final stop (Zsiga 2000: 311)

Figure 9 shows four examples of unreleased final stops, with variable amounts of voicing during the oral closure. This type of token represented the plurality (49%) of the data. Some tokens, such as 9A, showed no voicing during the oral closure and some, such as 9D, showed voicing throughout the oral closure. Most tokens, however, showed partial voicing of variable duration, as in 9B and C.



A. bought nine B. keep Matt

C. keep Matt D. keep Matt

Figure 9. Unreleased stops in Korean English, with variable voicing (Zsiga 2000: 318–320)

In addition, voicing of obstruents was found not only in pre-nasal position, but in all intersonorant environments. An example is shown in Figure 10. The phrase is "that ticket will" extracted from the target sentence "I hope that ticket will stop Nan from speeding," which was included in the dataset for the [p#n] sequence. The beginning words of the phrase, however, illustrate the pattern of voicing that was typical throughout the dataset. The initial /t/ and medial /k/ of "ticket" both show voicing for about half the closure duration, and the final /t/ of the word, preceding /w/, is fully voiced.

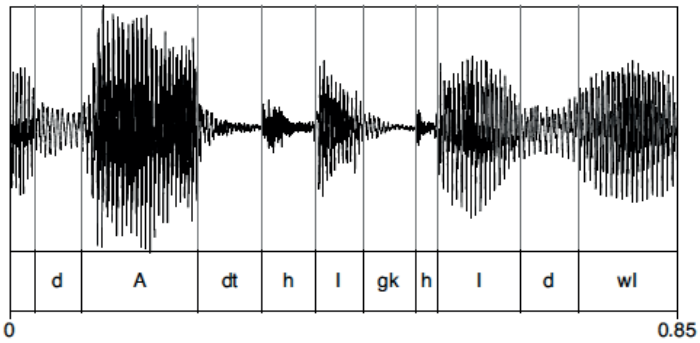


Figure 10. Waveform of "that ticket will" in Korean English, showing intersonorant voicing (Zsiga 2011: 320)

These patterns of intersonorant voicing show a close connection between words. The gradient and variable nature is indicative of a phonetic coarticulatory effect rather than phonological substitution. It is consistent with intersonorant voicing of lenis stops in native Korean, analyzed by S.-A. Jun (1995) as variable shortening and weakening of the laryngeal opening gestures. Some examples of intersonorant voicing in native Korean (from Silva 1992) are shown in (5).

(5) Intersonorant voicing in L1 Korean

/apeci -ka/ →	[abe̞iga]	"father-nominative"
/motun kulim/ →	[modun gulim]	"every picture"
/kulimul pota/ →	[kulimulboda]	"look at a picture"

The existence of gradient intersonorant voicing in the L2 English data indicates the transfer of coarticulatory routines (that is, patterns of gestural overlap) from native Korean to L2 English. The transfer is applying here not just within words, but across word boundaries. Overall, the voicing data supports an Articulatory Phonology account. What about the patterns of nasalization, which in native Korean appears to be category-changing phonology? Will connected speech phonological rules also transfer?

As was shown in Table 9, 23% of pre-nasal obstruents in this L2 dataset were pronounced as "full nasals," identical to an underlying nasal sequence. The formant patterns and amplitude are consistent with 100% nasality during the consonant sequence, and the durations of underlying and derived nasal sequences do not differ. For example, the mean duration of an [n#m] sequence derived from [t#m] was 149 ms, exactly the same as the mean for the an underlying [n#m] sequence. Figure 11 compares an underlying nasal sequence from the phrase "train Matt" to a derived nasal sequence from the phrase "ate mine." Both are pronounced as [n#m]. These "full nasals" are indicative of a categorical phonological substitution, not gestural overlap. The phonological rule of nasalization has transferred from L1 Korean to L2 English.

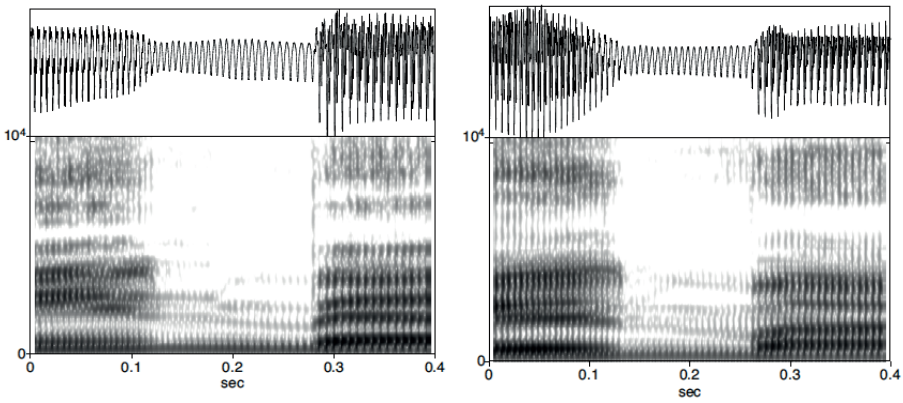


Figure 11. Underlying (left, "train Matt") and derived (right, "ate mine") [n#m] sequences in Korean English (Zsiga 2000: 312)

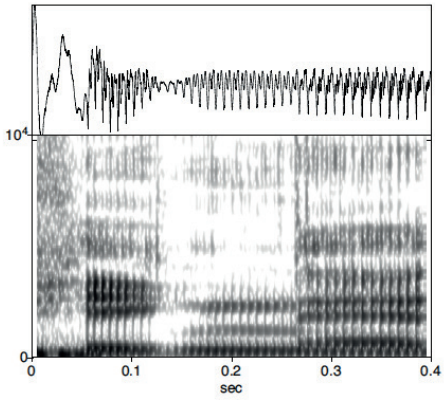


Figure 12. A spectrogram of the phrase "pick Nat" in Korean English showing partial nasalization (Zsiga 2000: 313)

Another 9% of the data, however, can be categorized as "partial nasals." These segments were transcribed as nasal consonants, but further acoustic analysis showed that the nasalization was partial and gradient. An example is shown in Figure 12, a spectrogram of the utterance "pick Nat" that was transcribed as "pi[ŋ n]at." The spectrogram shows that this is not a case of substitution, however. In this case, the consonant sequence begins with a very short oral period, transitioning quickly to nasality. These partial nasals are better analyzed in terms of gestural overlap rather than phonological substitution. They sound like nasal consonants, but the nasalization is gradient and variable.

Comparison of Figure 12 and Figure 7 shows similarity between the "partial nasal" realizations in L2 English and typical L1 English pronunciations. Both show close connection between words, and presumed gestural overlap at word boundaries. The difference is that during the oral part of the sequence, Korean native speakers show modal voicing, consistent with the pattern of their L1, while the American English speakers show glottalization.

These data, then, show evidence for two kinds of nasalization at word boundaries in Korean-English. In many cases, derived nasal sequences show formant patterns and amplitude consistent with nasality for 100% of closure. These are indistinguishable from underlying nasal sequences and are best analyzed as a categorical alternation. There are other sequences, however, that *sound* like nasals but show a gradient and variable pattern instead. This pattern is more like L1 English, but with modal voicing instead of glottalization. This pattern is best analyzed in terms of gestural overlap. There is evidence in this data for transfer of both L1 phonological substitution, for the full nasals, and of L1 gestural coordination, for the partial nasals and the voiced stops.

4. LEARNING FROM LEARNERS

In conclusion, what does the study of L2 English teach us about phonology? First, Word Integrity is at least partially wrong. Although one speaker did keep her words separate, she was in the minority. The majority of L2 English speakers in this experiment did "synchronize words," as evidenced by relatively high rates of nasalization and voicing across word boundaries. It remains unclear why the data in this experiment showed closer connections between words than previous experiments on Russian and Catalan speakers of English (Zsiga 2003, Cebrian 2000). It might be that the greater prevalence of connected speech assimilations is a fact particular to Korean or to Korean English. More data on different language pairs will help shed further light on this question.

Second, Articulatory Phonology is at least partially right. Most subjects showed gradient and variable intersonorant voicing, and many showed gradient and variable nasalization. These processes are consistent with the Articulatory Phonology analysis of gestural overlap at word boundaries.

However, there is also categorical assimilation across word boundaries. In native Korean, cross-word nasalization is consistently categorical. And in some cases, cross-word categorical nasalization transfers to L2 English. If nasalization was due to alternants listed in the lexicon, it would not transfer to English words. If it was due to gestural reorganization, it would not be categorical. Because it is in many cases categorical, and because it does transfer, these results show that Korean nasalization is a phonological alternation that is neither pre-specified in the lexicon nor the result of simple coarticulation. There is still a real phonology between the lexicon and speech production.

REFERENCES

Browman, C. and L. Goldstein. 1988. Some Notes on Syllable Structure in Articulatory Phonology. *Phonetica* 45, 140–155.

Browman, C. and L. Goldstein. 1990. Tiers in Articulatory Phonology, with Some Implications for Casual Speech. In J. Kingston and M. Beckman (eds.) *Papers in Laboratory phonology I: Between the Grammar and Physics of Speech*. New York: Cambridge University Press, 341–376.

Browman, C. and L. Goldstein. 1992. Articulatory Phonology: An Overview. *Phonetica* 49, 155–180.
Cebrian, J. 2000. Transferability and Productivity of L1 Rules in Catalan-English Interlanguage. *Studies in Second Language Acquisition* 22, 1–26.
Cohn, A. 1993. Nasalisation in English: Phonology or Phonetics. *Phonology* 10, 43–81.
Goldstein, L., D. Byrd, and E. Saltzman. 2006. The role of Vocal Tract Gestural Action Units in Understanding the Evolution of Phonology. In M. Arbib (ed.) *From Action to Language: The Mirror Neuron System*. Cambridge: Cambridge University Press, 215–249.
Hall, N. 2010. Articulatory Phonology. *Language and Linguistic Compass* 4/9, 818–830.
Huffman, M. 2005. Segmental and Prosodic Effects on Coda Glottalization. *Journal of Phonetics* 33, 335–362.
Jun, S.-A. 1995. Asymmetrical Prosodic Effects on the Laryngeal Gesture in Korean. In B. Connell and A. Arvaniti (eds.) *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*. Cambridge: Cambridge University Press, 235–253.
Pierrehumbert, J. 1995. Prosodic effects on glottal allophones. In O. Fujimura (ed.) *Vocal Fold Physiology 8*. San Diego: Singular Press, 39–60.
Silva, D. 1992. *The phonetics and phonology of stop lenition in Korean*. Doctoral dissertation, Cornell University.
Weinberger, S. 1994. Please speak faster, I can't understand you: On the L2 transfer of fast speech processes. Paper presented at the Second Language Research Forum, Montreal.
Zsiga, E. 1995. An Acoustic and Electropalatographic Study of Lexical and Post-lexical Palatalization in American English. In B. Connell & A. Arvaniti (eds.) *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*. Cambridge: Cambridge University Press, 282–302.
Zsiga, E. 2000. Phonetic Alignment Constraints: Consonant Overlap and Palatalization in English and Russian. *Journal of Phonetics* 28, 69–102.
Zsiga, E. 2003. Relearning Consonant Timing. *Studies in Second Language Acquisition* 25, 399–432.
Zsiga, E. 2011. External Sandhi in a Second Language: The Phonetics and Phonology of Obstruent Nasalization in Korean-accented English. *Language* 87, 289–345.

SUMMARY

THE PHONETICS AND PHONOLOGY OF ENGLISH CASUAL SPEECH: LEARNING FROM L2 LEARNERS

This paper examines processes of "connected" or "casual" speech in second language pronunciation, focusing on the speech of Korean learners of English. The paper begins with the point of view of Articulatory Phonology, which argues that many assimilations and deletions in casual speech are the result of overlap between articulatory gestures. Examples from English and Russian illustrate gestural overlap. Further examples are provided from a more detailed phonetic study of processes of nasalization and voicing assimilation in Korean and Korean-accented English (Zsiga 2011). The Korean-

English data show evidence of gradient gestural overlap in voicing assimilation, and in some instances of partial nasal assimilation, supporting the Articulatory Phonology approach. Many instances of categorical nasal substitution were also found, however. It is argued that a more traditional phonological feature-changing analysis better accounts for the categorical changes. Both Articulatory Phonology and traditional feature-based phonology are required to account for the full set of data.

KEYWORDS: casual speech, second language learning, Phonology, Articulatory Phonology, Korean, accented English, nasal assimilation.

ARTICLE INFO:
Original research article
Received: November 11, 2018
Revised: February 11, 2019
Accepted: February 12, 2019

UDC: 81'243
811.111'355(497.11)
DOI: https://doi.org/10.18485/philologia.2018.16.16.2

FORMANT MEASUREMENTS OF SERBIAN SPEAKERS' ENGLISH VOWELS

ANDREJ BJELAKOVIĆ[1]
English Department
Faculty of Philology
University of Belgrade
Belgrade, Serbia

U radu se predstavljaju rezultati akustičkog ispitivanja produkcije vokala engleskog jezika u govoru izvornih govornika srpskog jezika, studenata Katedre za anglistiku Filološkog fakulteta Univerziteta u Beogradu. Broj učesnika ispitivanja bio je 26 (13 studenata prve godine i 13 studenata završnih godina). Mereni su F1 i F2 naglašenih vokala — 11 monoftonga (KIT, DRESS, TRAP, FOOT, STRUT, LOT, FLEECE, PALM, GOOSE, THOUGHT i NURSE) i 4 diftonga (GOAT, PRICE, MOUTH i FACE). Takođe su mereni vokali u L1 ispitanika, koje smo poredili sa vokalima u njihovom L2, to jest engleskom. Ispitanici su snimljeni dok čitaju izabrane odlomke vesti Bi-Bi-Sija i dve kratke priče na srpskom. Broj analiziranih primera je 7534 na engleskom (oko 305 vokala po govorniku), i 4266 na srpskom. Rezultati pokazuju da učenici engleskog jezika čiji je maternji jezik srpski, na nivou znanja naših ispitanika, načelno ne zamenjuju vokale L2 vokalima svog L1; s druge strane, ne može se reći da dosežu vrednosti karakteristične za izvorne govornike, već kompromisne vrednosti. Izuzeci, kada je reč o zameni jesu vokal DRESS, i bar za neke govornike, vokal TRAP (oba su supstituisana srpskim kratkim /e/). Vokali kod kojih se javila najveća razlika između dve podgrupe ispitanika bili su GOAT, GOOSE, MOUTH, PRICE, i u manjoj meri THOUGHT, pri čemu su stariji, iskusniji studenti proizvodili vrednosti više nalik onima kod izvornih govornika.

Ključne reči: kvalitet vokala, usvajanje L2, L2 izgovor, EFL izgovor, vokali engleskog, frekvencija formanata, srpski.

1 Contact information: b.andrej@fil.bg.ac.rs

1. INTRODUCTION

The aim of the present paper is to add to the store of acoustic data pertaining to the vowels of English as produced by native speakers of Serbian. These data can shed light on how L1 phonology affects learners' acquisition of L2 phonology. Previous studies which acoustically investigated Serbian speakers' production of English include Marković (2007) and Marković (2012) (the results of which have also been partially presented in Marković 2009a, Marković 2009b, and Marković & Jakovljević 2013), Mlinar (2011), Paunović (2002) and Paunović (2011), Sudimac (2016), Marković & Jakovljević (2016), Dančetović & Nešić (2017); a subset of these involve Serbian speakers who had been living in an English speaking country prior to participating in the experiment, such as Krebs-Lazendic 2008, Krebs-Lazendic & Best 2007, Krebs-Lazendic & Best 2013, Čubrović 2016 and Čubrović 2017.

Our approach will consist of comparing the Serbian speakers' English vowels on the one hand with their L1 vowels (vowels of contemporary Belgrade Serbian), which we also analysed, and on the other hand with vowels of English native speakers (as presented in Bjelaković 2017).

We shall look at the results in the light of, on the one hand, Flege's Speech Learning Model (SLM) (Flege 1988, Flege 1990, Flege 1995, Flege 2005), especially hypotheses 3–7 (Flege 1995: 239), which together predict that L2 phonemes that are phonetically more dissimilar to L1 phonemes will more readily develop as new categories for learners; on the other hand, we will test the hypothesis that learners' experience will have effect on their L2 production (Bohn & Flege 1992, Flege et al. 1997, Munro & Derwing 2008, Derwing et al. 2007). This hypothesis has been confirmed in the papers cited, however, given the relatively small gap between our two groups of participants (only 3 years of learning and speaking English), we want to see if any measurable effect of this gap can be detected. Furthermore, in relation to that, we are interested to see which sounds will exhibit this effect, as Bohn and Flege's (1992) hypothesis predicts that groups with different experience levels will not exhibit differences when producing 'similar' sounds, whereas they will exhibit them when producing 'new' sounds.

2. METHOD

The participants in the present study (n=26) were all students of the English Department at the Faculty of Philology, University of Belgrade[2]. They were divided into two groups of equal size (each group had 7 female and 6 male participants), with the first group comprising freshmen (S1–S13), and the second group comprising fourth year and MA students (S14–S26). All participants grew up in Belgrade and are consequently speakers of the same L1 variety. All participants had been learning English formally between 10

2 Participant selection involved a survey with questions "Which do you prefer listening to?" and "Which do you prefer using?", the aim of which was to ensure participants preferred British as opposed to American pronunciation. This was done due to the fact that reference native speaker vowels from Bjelaković (2017) are vowels of Standard British English.

and 15 years, and none of them had ever lived in an English-speaking country (though some of the participants had visited English speaking countries briefly, on holiday).

The participants were recorded while performing reading tasks, using the Tascam DR-100mkII digital recorder, placed around 40 cm away from the participant's mouth.

The first reading task involved two short stories in Serbian (327 words and 263 words, respectively). The number of analysed vowel tokens ranged between 159 and 169 per speaker, yielding a total of 4266 analysed tokens. The number of analysed vowels for each vocalic category was as follows[3]: 18 /a/, 15 /a:/, 27 /e/, 18 /e:/, 13 /i/, 12 /i:/, 31 /o/, 8 /o:/, 15 /u/, 12 /u:/.

The second reading task involved paragraphs of BBC news copy, taken from Bjelaković (2017) (a total of 1923 words). The number of analysed vowel tokens was around 307 per speaker, with a total of 7534 analysed tokens. The following English vowels were analysed (the number of tokens is in parentheses): DRESS[4] (28), FACE (21), FLEECE (20), FOOT (14), GOAT (17), GOOSE (17), KIT (22), LOT (21), MOUTH (16), NURSE (22), PALM (19), PRICE (23), STRUT (17), THOUGHT (26) and TRAP (24).

Only stressed vowels were analysed. Regarding the phonetic environment, vowels that were either before or after /w/, /r/, /j/ or /l/ were completely avoided (see Deterding 1997: 49), as these approximants would most likely have noticeable coarticulatory effects. Other sonorants (i.e. nasals or other vowels) were also avoided whenever possible. The above conditions were somewhat relaxed with the FOOT vowel, due to its relative infrequency.

All tokens were analysed manually with Praat, v. 5.4 (Boersma and Weenink 2014), using the get formant function for F1 and F2. As is customary (Boersma 2013: 395), the settings with the cut-off at 5,500 Hz, were used for tracking all female speakers' formants, and the cut-off was lowered to 5,000 Hz for the analysis of the recordings from the male speakers.

A steady-state area around the temporal midpoint of the vowel was where measurements were taken for monophthongs. In cases where no steady state was present, as with some short monophthong tokens, measurements were made at the exact midpoint of the vowel's duration. Regarding diphthongs, to avoid the most obvious consonant coarticulation effects, measurements of the onset part were typically made after the first 12–16 per cent of the vowel's duration, while the glide was analysed by taking measurements at around 79–85 per cent of the vowel's duration, i.e. close to the endpoint of the vowel. The exact point of measurement was determined taking into account the central tendency of the vowel in question and the maximal value reached (e.g. for MOUTH onsets both F1 and F2 were measured at the point of highest F1 reached) (see Labov et al. 2006: 38).

The same method for analysing vowel formants was used in Bjelaković (2017), which analysed the speech of BBC newsreaders, the results of which will serve as reference values throughout the present paper.

3 Previous studies (e.g. Lehiste & Ivić 1986: 67) have found that vowel length in Serbian has an effect on vowel quality (in non-high vowels), which is why we analyse the Serbian vocalic system as having five short and five long vowels. This approach was originally proposed by Jakobson (1937 [1962]).

4 As per usual, lexical set keywords originally proposed in Wells (1982) are used.

3. RESULTS

For each set of vowels, the following pattern will be used: the first figure will show Serbian speakers' L1 and L2 vowels, while the next figure will show Serbian speakers' English vowels and BBC newsreaders' vowels (adapted from Bjelaković 2017). As mentioned before, Serbian speakers comprised two groups: a less experienced first-year student group, and a more experienced fourth year/fifth year group; however, the two will be conflated in most figures, except for those vowels where the difference between two groups is noteworthy. Figures will show mean values normalized according to Lobanov (1971) and will be created using NORM (Thomas & Kendall 2007); ellipses will represent one standard deviation. Finally, a table with raw Hertz values will also be provided for each vowel.

3.1 MONOPHTHONGS

3.1.1 FLEECE[5] and KIT vs. Serbian /i/

Figure 1 displays the Serbian speakers' L1 vowels (/i/ and /i:/) and their L2 vowels (KIT and FLEECE). The Serbian vowels show little overlap with the English vowels (especially with KIT), and occupy the space between them. In other words, Serbian speakers' FLEECE vowel is more peripheral than their Serbian /i/~/i:/, while their KIT vowel is more centralised.

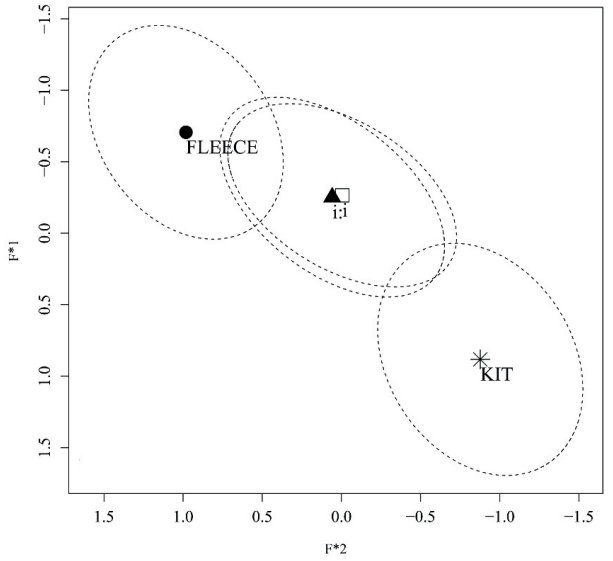


Figure 1. Mean formant values of Serbian speakers' L1 and L2 vowels, normalised according to Lobanov (1971)

5 Vowels FLEECE and GOOSE are treated as monophthongs in this analysis. In other words, the slight degree of diphthongisation present was not analysed, and measurements of F1 and F2 were made at one point only.

The following figure (Figure 2) shows BBC newsreaders' vowels (labelled FLEECEn and KITn) alongside Serbian speakers' KIT and FLEECE. As is apparent, FLEECE and FLEECEn overlap, with FLEECEn having a smaller ellipse, indicating less variation. On the other hand, the native KIT value (KITn), in addition to also having a smaller ellipse, is more centralised. This indicates that not all Serbian speakers have adequately acquired a native-like quality of KIT.

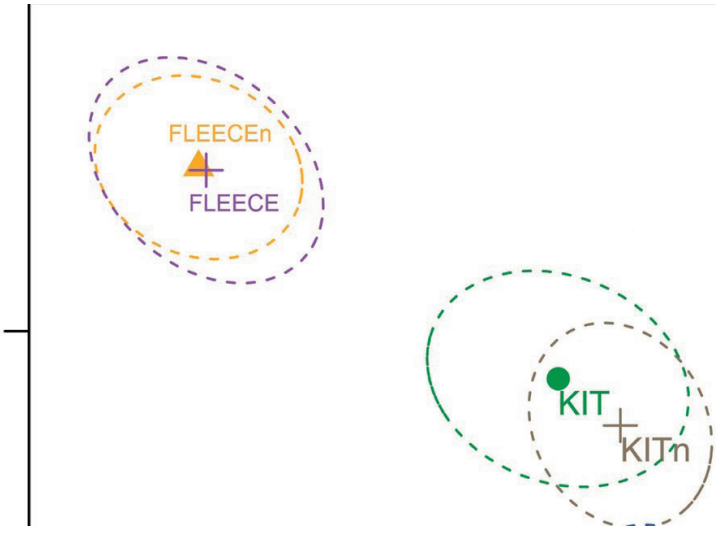


Figure 2. Mean formant values of Serbian speakers' (KIT, FLEECE) and BBC newsreaders' (KITn, FLEECEn) English vowels, normalised according to Lobanov (1971)

However, looking at individual vowel charts (not shown here) we see that only three Serbian speakers (two younger (S2 and S5) and one older (S25)) display a certain degree of overlap of KIT and FLEECE.

On the whole, the older and the younger student group displayed very similar results when it came to these two vowels.

	FLEECE		KIT		/i:/		/i/	
BBCM	290	2367	394	1830				
BBCF	348	2623	458	2073				
M	329	2199	413	1825	355	2039	347	2038
F	391	2633	476	2120	413	2355	420	2318

Table 1. Raw mean F1 and F2 frequencies of BBC male and female newsreaders (BBCM, BBCF), and Serbian speakers (M – male, F – female).

3.1.2 DRESS and TRAP vs. Serbian /e/

It is generally acknowledged that Serbian learners of English have difficulties acquiring the TRAP/DRESS opposition (Marković 2007, 2009, 2012; Paunović 2011, Marković & Jakovljević 2016). The same is true of native speakers of other languages, like Spanish or German (Flege et al. 1997).

In the figure below (Figure 3) it can be seen that the Serbian short /e/ largely overlaps with the Serbian production of the DRESS vowel. This tallies well with the predictions of Flege's Speech Learning Model (Flege 1995), as well as the results in Marković (2012: 105). Specifically, DRESS is categorised as a 'similar' vowel, and thus assimilated to the L1 category of /e/. On the other hand, the Serbian /e:/ is closer and fronter, and a clear separation is maintained. As regards TRAP Serbian participants show a degree of overlap with DRESS.

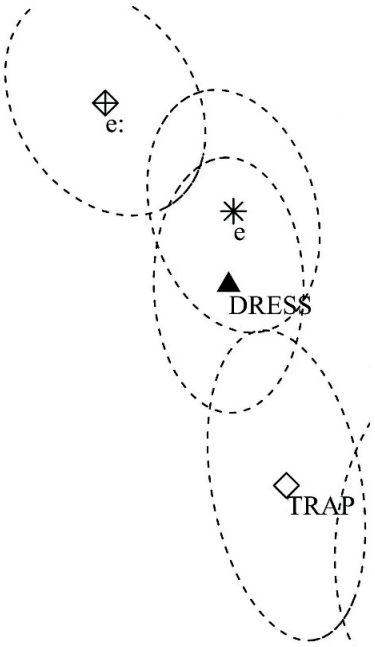


Figure 3. Mean formant values of Serbian speakers' L1 and L2 vowels, normalised according to Lobanov (1971)

The native TRAP/DRESS configuration is shown against that of the Serbian participants' in Figure 4. It is apparent that the native DRESS vowel (DRESSn) is somewhat closer and fronter than the Serbian learners' DRESS vowel, while the native TRAP vowel (TRAPn) is more open than the Serbian learners' TRAP vowel. Consequently, the native TRAP and DRESS display quite a clear separation, which is not the case with those produced by the Serbian participants.

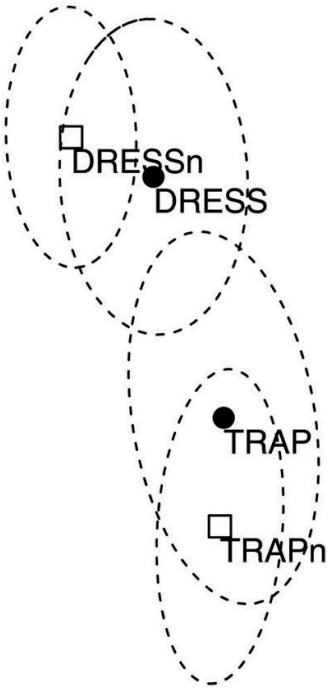


Figure 4. Mean formant values of Serbian speakers' (DRESS, TRAP) and BBC newsreaders' (DRESSn, TRAPn) English vowels, normalised according to Lobanov (1971)

When it comes to mean values, especially for TRAP, the older and the younger group, as with FLEECE and KIT, showed very similar results (the statistical analysis in Section 4 below confirms this). However, looking at individual vowel charts (not shown here), we can divide the Serbian participants into three groups: ellipses of TRAP and DRESS show a slight degree of overlap (four younger participants, S1, S4, S7 and S10, and five older participants S17, S19, S20, S22 and S24); ellipses of TRAP and DRESS show a large degree of overlap (two younger participants, S2 and S5, and three older participants, S16, S23 and S25); there is no overlap of TRAP and DRESS ellipses (seven younger participants S3, S6, S8, S9, S11, S12 and S13, and five older participants S14, S15, S18, S21 and S26) (see section 4).

	DRESS		TRAP		/e:/		/e/	
BBCM	544	1722	699	1546				
BBCF	615	1913	841	1665				
M	567	1582	673	1504	494	1769	525	1593
F	714	1824	830	1703	582	2046	676	1801

Table 2. Raw mean F1 and F2 frequencies of BBC newsreaders and Serbian speakers

3.1.3 PALM and STRUT vs. Serbian /a/

In this section we will be comparing the English PALM vowel with the Serbian /a:/ and the English STRUT vowel with the Serbian /a/. Figure 5 displays the Serbian participants' English and Serbian vowels.

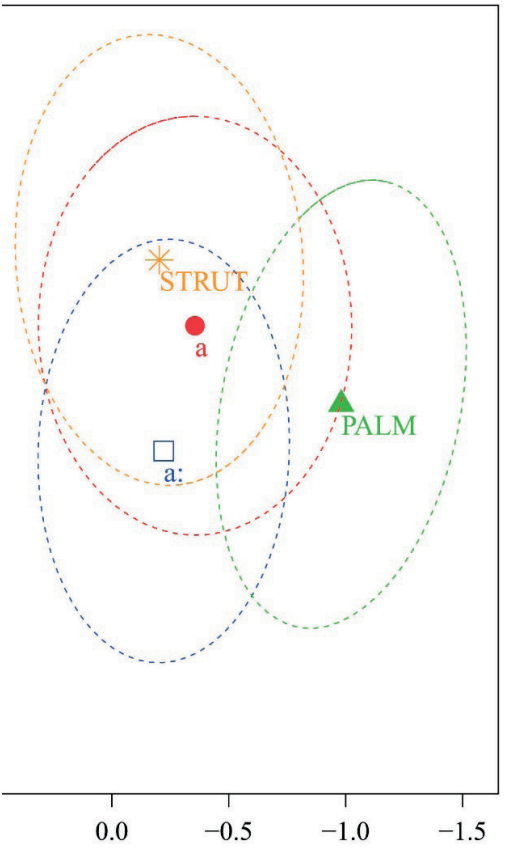


Figure 5. Mean formant values of Serbian speakers' English and Serbian vowels, normalised according to Lobanov (1971)

It is apparent that there is a large degree of overlap between the Serbian /a/ and the Serbian speakers' STRUT vowel, with the latter being somewhat more centralised. On the other hand, the PALM vowel does not overlap with the Serbian /a:/ quite as much, and is more retracted, as well as somewhat raised.

In Figure 6 we compare the Serbian speakers' PALM and STRUT with the BBC newsreaders' PALM and STRUT ('PALMn' and 'STRUTn').

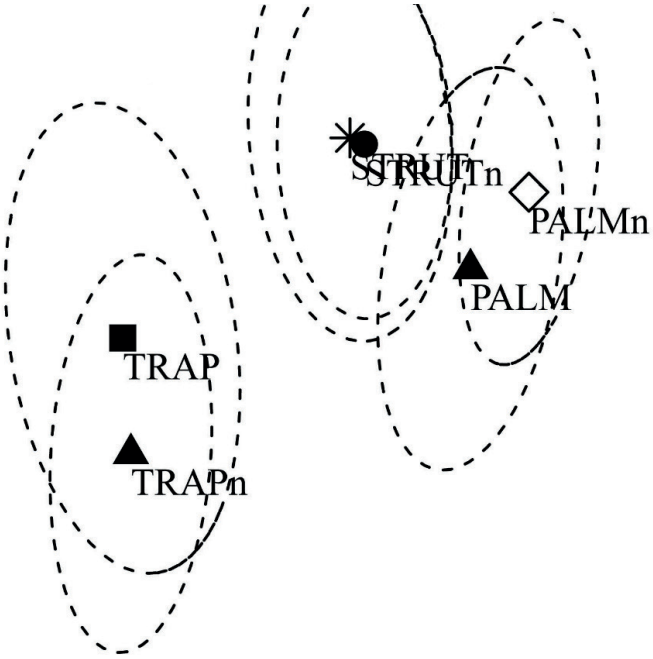


Figure 6. Mean formant values of Serbian speakers' (PALM, STRUT, TRAP) and BBC newsreaders' (PALMn, STRUTn, TRAPn) English vowels, normalised according to Lobanov (1971) (ellipses of other vowels, such as LOT, have been removed for the sake of clarity).

As before, the native speakers' ellipses are somewhat smaller. The STRUT vowel is remarkably similar (the mean values of STRUT and STRUTn are virtually the same). The mean value of the native speakers' PALM vowel ('PALMn'), however, is more retracted and raised. In other words, Serbian participants seem to produce an intermediate quality, one between the Serbian /a:/ and the native quality of PALM.

	STRUT		PALM		/a:/		/a/	
BBCM	611	1264	625	1120				
BBCF	697	1418	719	1191				
M	623	1263	674	1165	653	1266	619	1243
F	736	1438	782	1271	818	1432	781	1409

Table 3. Raw mean F1 and F2 frequencies of BBC newsreaders and Serbian speakers

3.1.4 LOT and THOUGHT vs. Serbian /o/

It is worth noting at the outset that a minority of Serbian participants tended to use a quality for LOT that was more like that found in North American English, i.e. more open and less rounded, especially in frequent words such as *not*, *body*, *God*, *job* and *shot*. These tokens were also analysed, primarily because it was difficult to draw the line and decide which tokens were to be inadmissible.

In Figure 7 we can see that the difference between Serbian /o/ and /o:/ is smaller than the difference between their /e/ and /e:/ (see Figure 3 above), and that this Serbian vowel is situated between the Serbian speakers' LOT, which is noticeably more open, and their THOUGHT, which is decidedly closer and retracted.

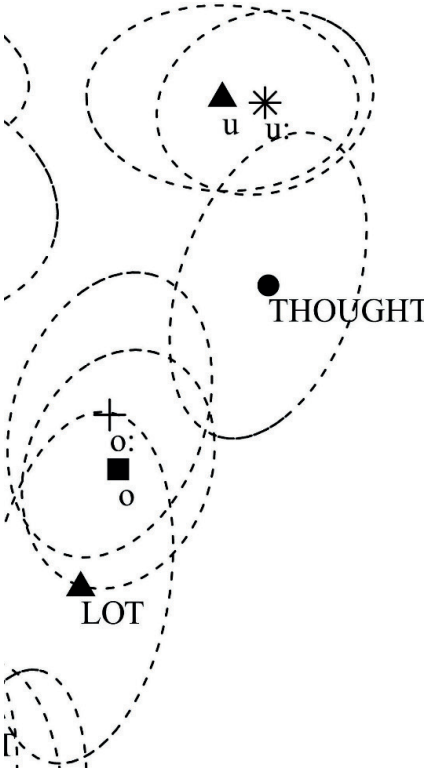


Figure 7. Mean formant values of Serbian speakers' English and Serbian vowels, normalised according to Lobanov (1971)

However, looking at the individual vowel charts, not shown here, we can see that the Serbian speakers' LOT vowel differs from their THOUGHT in one respect. Namely the LOT vowel shows much more individual variation, so that some speakers display the configuration as the one in Figure 7, while others display a great degree of overlap between their LOT vowel and the Serbian /o/ and/or /o:/ (two younger speakers, S1 and

S5, and three older speakers, S17, S18 and S22). On the other hand, all speakers showed a clear separation of their THOUGHT vowel from the Serbian vowels, indicating that this category is better acquired.

Figure 8 shows the Serbian speakers' and the BBC newsreaders' English vowels (LOT, THOUGHT, and LOTn, THOUGHTn, respectively). The native speakers' ellipses are yet again smaller, and this is especially true of THOUGHTn (the centre of which is a triangle in Figure 8). Native speakers had closer qualities of both LOT and THOUGHT.

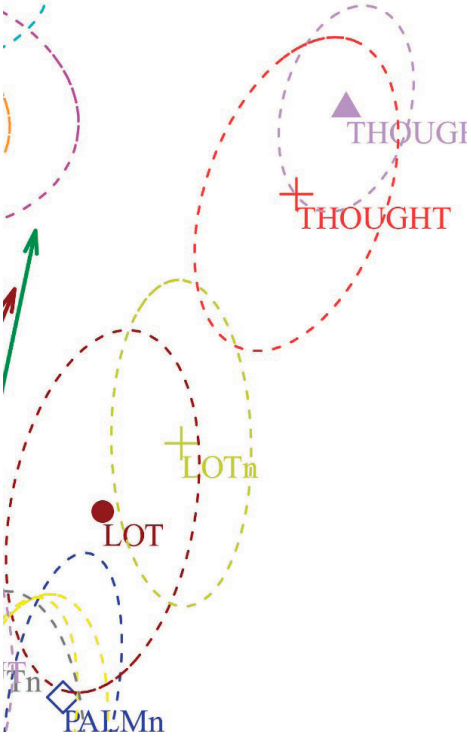


Figure 8. Mean formant values of Serbian speakers' (LOT, THOUGHT) and BBC newsreaders' (LOTn, THOUGHTn) English vowels, normalised according to Lobanov (1971)

In Figure 9 we split the Serbian participants into two subgroups, the younger, first-year student group (LOTml, THOUGHTml), and the older, more experienced group (LOTst, THOUGHTst), in addition to the BBC newsreaders (LOTn, THOUGHTn). There we can see that for each of the vowels the older students' vowel qualities were closer to those of the native speakers.

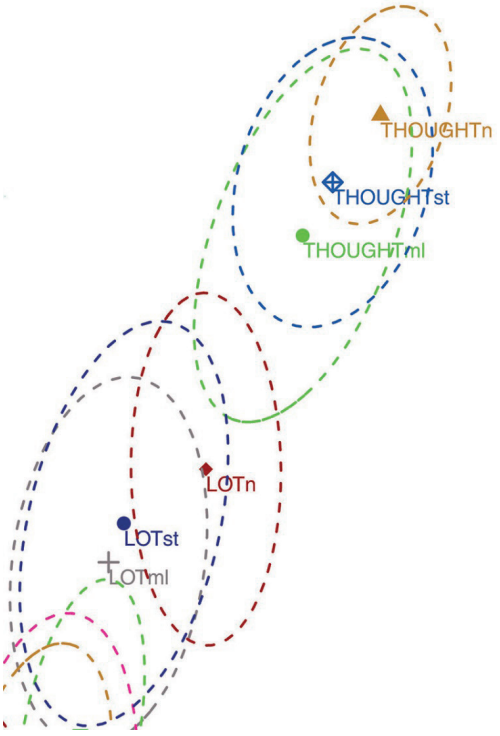


Figure 9. Mean formant values of less experienced (LOTml, THOUGHTml) and more experienced (LOTst, THOUGHTst) Serbian speakers' English vowels, as well as BBC newsreaders' (LOTn, THOUGHTn) vowels, normalised according to Lobanov (1971)

	LOT		THOUGHT		/o:/		/o/	
BBCM	547	959	407	750				
BBCF	577	1039	419	821				
M	574.2	1072	472.3	819.4	497	1023	514	1014
F	665.9	1181	506.6	922.5	594	1155	626	1139

Table 4. Raw mean F1 and F2 frequencies of BBC newsreaders and Serbian speakers

3.1.4 FOOT AND GOOSE VS. SERBIAN /U/

The Serbian /u/ is a back vowel, whereas the English GOOSE vowel, although once back, has been fronting for much of the 20th century (Wells 1997), so that now, at the beginning of the 21st century this is a central vowel, with a tendency to become a front vowel for some speakers (Windsor Lewis 1995, Cruttenden 2014: 133). The FOOT vowel seems to have been following this fronting, albeit to a lesser degree.

In Figure 10 we can see that Serbian participants do not, on the whole, exhibit any overlap between their Serbian back /u/ and their English FOOT and GOOSE vowels. The latter two, however, do partly overlap, indicating an imperfect acquisition of the FOOT vowel's centralisation. Also apparent is that their GOOSE ellipse is remarkably wide, indicating a great deal of variation.

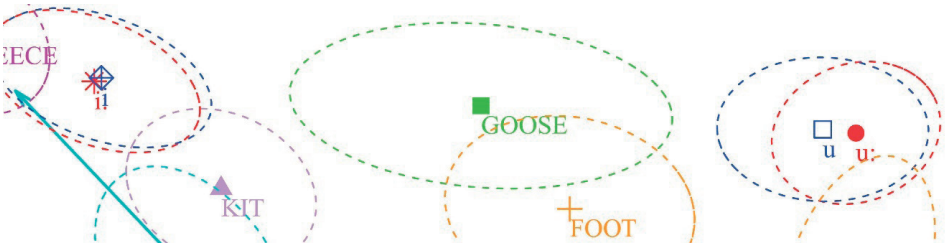


Figure 10. Mean formant values of Serbian speakers' English and Serbian vowels, normalised according to Lobanov (1971)

Looking at individual charts (not shown here) we can see that the Serbian and English vowel ellipses barely overlap for two younger (S2 and S7) and two older speakers (S20 and S25), whereas they overlap considerably for two younger (S5 and S11) and one older speaker (S16).

Figure 11 shows Serbian speakers' and BBC newsreaders' English vowels (FOOT, GOOSE, and FOOTn, GOOSEn, respectively). Apparent is the fact that the native speakers' GOOSE vowel (GOOSEn) has a much smaller ellipse and that its mean F2 value is much higher, that is to say the native speakers' GOOSE vowel is consistently fronter. In addition to this, the native speakers' FOOT and GOOSE show very clear separation, with GOOSE being not only fronter but more peripheral, as expected.

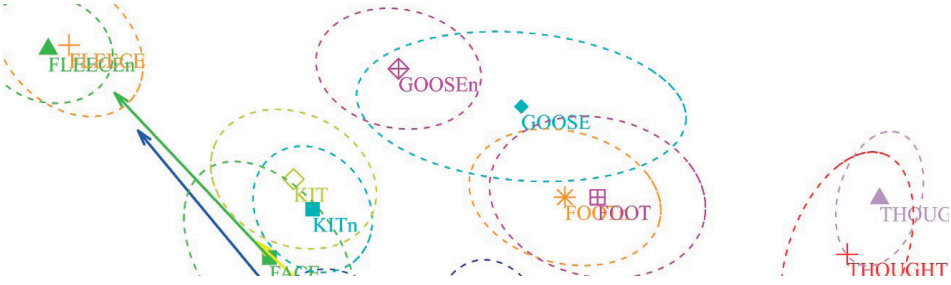


Figure 11. Mean formant values of Serbian speakers' (FOOT, GOOSE) and BBC newsreaders' (FOOTn, GOOSEn) English vowels, normalised according to Lobanov (1971)

Next, in Figure 12, we will split the Serbian participants into two subgroups, the younger, first-year student group (FOOTml, GOOSEml), and the older, more experienced group (FOOTst, GOOSEst).

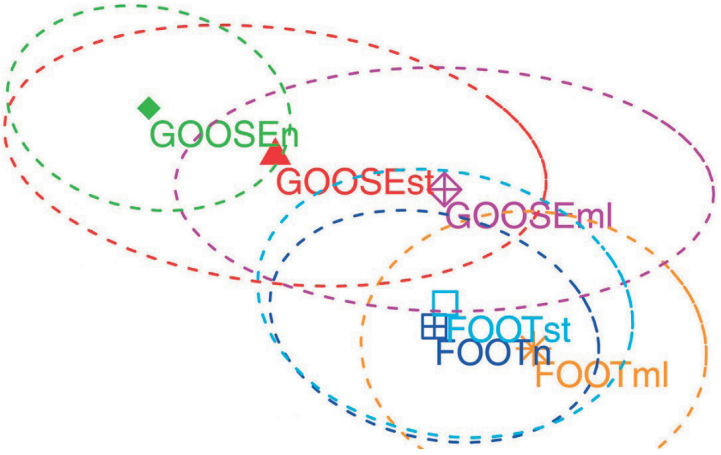


Figure 12. Mean formant values of less experienced (FOOTml, GOOSEml) and more experienced (FOOTst, GOOSEst) Serbian speakers' English vowels, as well as BBC newsreaders' (FOOTn, GOOSEn) vowels, normalised according to Lobanov (1971) (ellipses of other vowels, such as KIT, have been removed for the sake of clarity)

Figure 12 shows that older, experienced students exhibit FOOT qualities very close to those of the native speakers, while younger students' FOOT is more retracted. However, the difference between the two groups is larger when it comes to the GOOSE vowel. Namely, the mean quality of the older students' GOOSE (GOOSEst) is halfway between the rather front qualities of native speakers, and the backer qualities of younger students. Still, the older students' ellipse is not smaller than that of the younger speakers, indicating that they too exhibited a large degree of variation.

Looking at individual charts (not shown here) we can see that the following Serbian speakers reached native-like GOOSE qualities: two younger participants (S8 and S12), and four older participants (S14, S19, S20 and S21).

	GOOSE		FOOT		/u:/		/u/	
BBCM	317	1681	391	1349				
BBCF	343	1849	448	1490				
M	365	1373	420	1267	386	859	379	907
F	430	1683	491	1483	440	903	444	965

Table 5. Raw mean F1 and F2 frequencies of BBC newsreaders and Serbian speakers

3.1.5 NURSE

Vowels of Serbian that are closest in terms of quality to the mid-central RP NURSE are the short /e/ (which is front, but of appropriate height), and short /a/ (which is central but more open than NURSE).

Figure 13 shows that the mean value of Serbian speakers' NURSE was somewhat more open and front that the native speakers' NURSE. Individual vowel charts show that for native speakers the one standard deviation ellipse of NURSE does not overlap at all with any other monophthong; on the other hand, for the majority of Serbian speakers (14 out of 26) there is at least a slight overlap between NURSE and the Serbian short /e/ (and for some speakers the DRESS vowel as well).

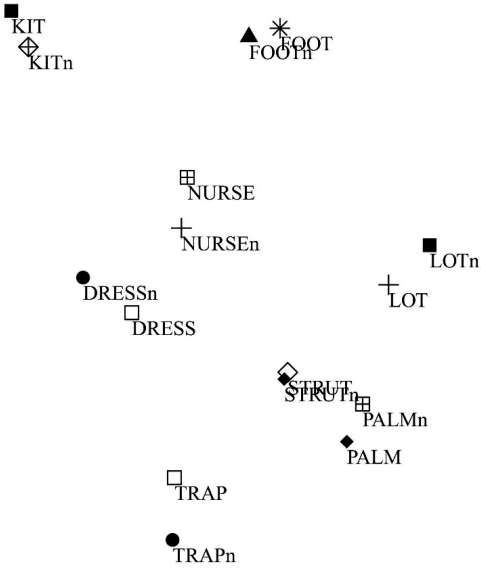


Figure 13. Mean formant values of Serbian speakers' and BBC newsreaders' English vowels, normalised according to Lobanov (1971)

	NURSE	
BBCM	505	1489
BBCF	597	1684
M	504	1434
F	605	1719

Table 6. Raw mean F1 and F2 frequencies of BBC newsreaders and Serbian speakers

3.2 DIPHTHONGS

3.2.1 PRICE and MOUTH

We will begin by comparing the lower section of the vowel space of the BBC newsreaders (Figure 14, top) and Serbian participants (Figure 14, bottom).

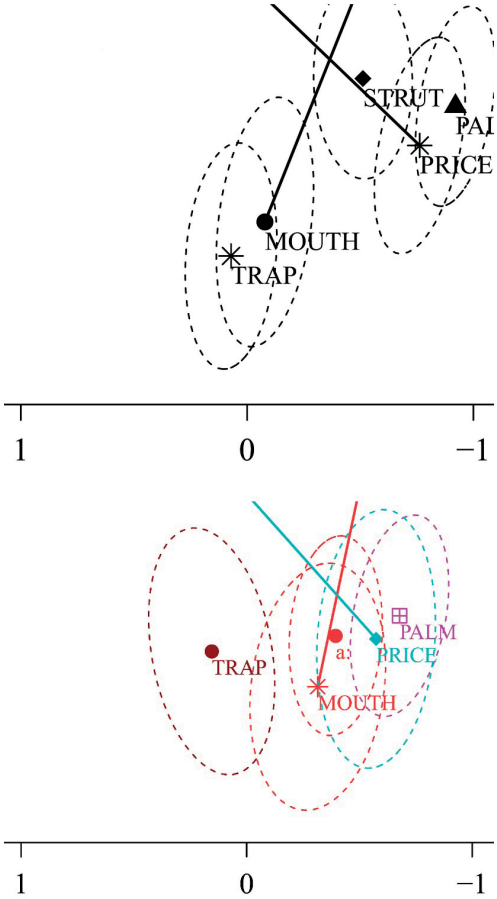


Figure 14. The lower section of the vowel space of the BBC newsreaders (top) and Serbian participants (bottom); the mean formant values were normalised according to Lobanov (1971) and the ellipses are one standard deviation (ellipses of other vowels, such as DRESS and the Serbian STRUT and short /a/ etc., have been removed for the sake of clarity).

In terms of similarities we can note that both native speakers' and Serbian speakers' onset of MOUTH is fronter than their onset of PRICE, or at least the mean values thereof; however, the onset of MOUTH is not as front as TRAP, nor is the onset of PRICE as back as PALM. The Serbian long /a:/ is between these two onsets, for our Serbian speakers.

On the other hand the one standard deviation ellipses are much larger for the Serbian speakers, and overlap with one another. The reason for this is that for many of the Serbian participants the onsets of these two diphthongs are very close to one another.

Looking at the individual charts not shown here we can see that the PRICE and MOUTH ellipses do not overlap for the following Serbian participants: one younger (S12) and eight older (S14, S17, S18, S19, S21, S22, S24 and S26) (these participants are also the ones whose PRICE and MOUTH display the least degree of overlap with the Serbian long /a:/).

In Figure 15 we separate the Serbian participants into two subgroups, the younger, first-year student group ('ml'), and the older, more experienced group ('st'). There we can see that the onset of MOUTH for older participants is halfway between the onset of native speakers and that of the younger, first-year participants. Similar is true for PRICE, at least when it comes to F2 values, i.e. the horizontal dimension. In other words, the distance between the onsets of PRICE and MOUTH is the greatest for native speakers, followed by the more experienced students, with the first-year students displaying the shortest distance between the two.

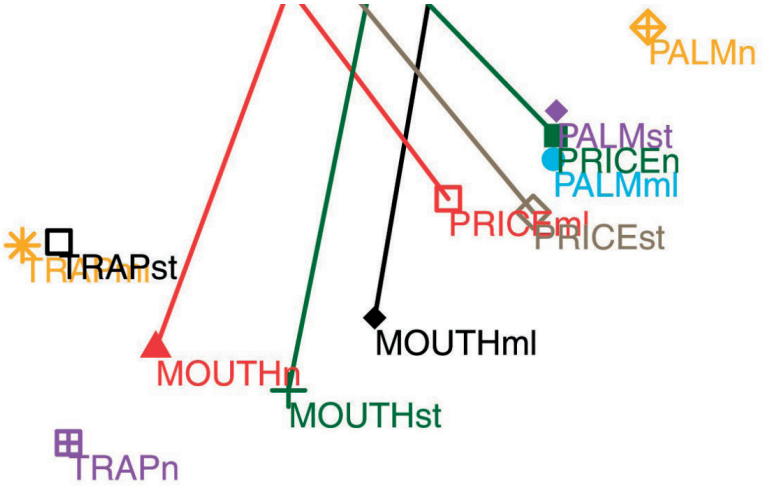


Figure 15. Mean formant values of less experienced (TRAPml, MOUTHml, PRICEml, PALMml) and more experienced (TRAPst, MOUTHst, PRICEst, PALMst) Serbian speakers' English vowels, as well as BBC newsreaders' (TRAPn, MOUTHn, PRICEn, PALMn) vowels, normalised according to Lobanov (1971)

	PRICE ons.		PRICE glide	
BBCM	635	1177	396	1946
BBCF	769	1277	496	2178
M	686	1200	485	1702
F	793	1332	560	2033
	MOUTH ons.		MOUTH glide	
BBCM	659	1440	435	1118
BBCF	824	1622	539	1303
M	699	1293	474	1158
F	844	1478	539	1272

Table 7. Raw mean F1 and F2 frequencies of BBC newsreaders and Serbian speakers

3.2.1 FACE

In Figure 16 we display the BBC newsreaders' front mid portion of the vowel space. As is apparent, the onset of FACE overlaps with the DRESS vowel. Its glide ends between their KIT and FLEECE monophthongs.

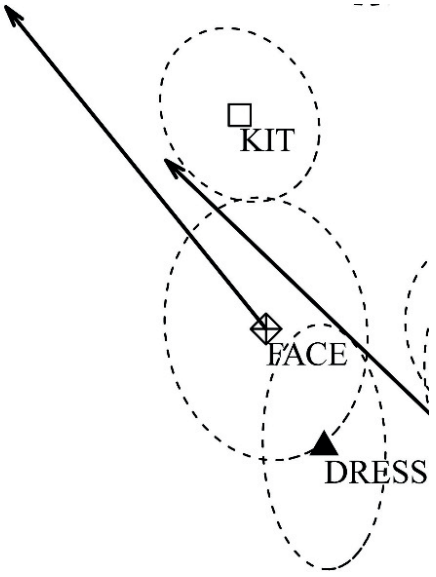


Figure 16. BBC newsreaders' front mid portion of the vowel space; mean formant values were normalised according to Lobanov (1971); ellipses are one standard deviation.

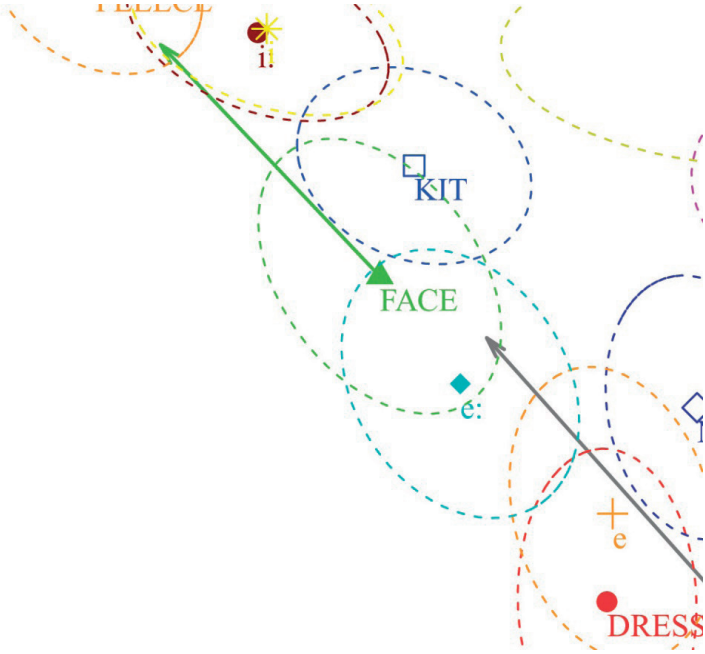


Figure 17. Serbian participants' front mid portion of the vowel space; mean formant values were normalised according to Lobanov (1971); ellipses are one standard deviation.

On the other hand, in Figure 17, which displays the Serbian participants' front mid portion of the vowel space, we can see that their FACE overlaps neither with their DRESS nor with their Serbian short /e/, but rather with their KIT and the Serbian long /e:/. Its glide enters the FLEECE and Serbian long /i:/ territories. Comparing Figures 16 and 17 it can also be noticed that the Serbian participants' FACE is a narrower diphthong (i.e. there is less distance between its onset and the end of its glide).

The same can be seen in Figure 18, which shows both the Serbian speakers' and the BBC newsreaders' English vowels. Specifically, it is apparent that DRESSn i FACEn are fairly close to one another, whereas the Serbian speakers' DRESS is more open and their FACE onset closer.

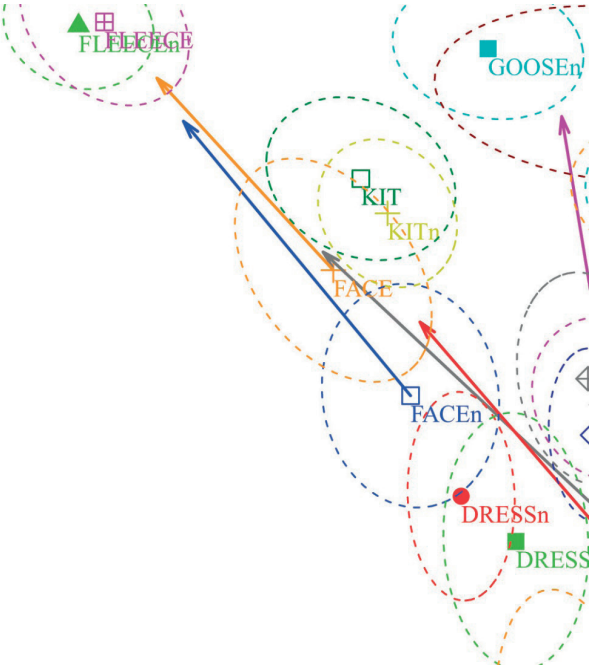


Figure 18. Mean formant values of Serbian speakers' (DRESS, FACE, KIT) and BBC newsreaders' (DRESSn, FACEn, KITn) English vowels, normalised according to Lobanov (1971)

In charts with separated younger and older Serbian speakers (not shown here) it can be seen that the two subgroups display very similar FACE qualities. It seems, then, that Serbian learners of English tend to substitute the onset of this diphthong with their L1 long /e:/ (the fact that their FACE onset tends to be closer still would be due to the coarticulation with the closing glide).

	FACE ons.		FACE glide	
BBCM	493	1761	341	2204
BBCF	565	2067	404	2416
M	463	1854	366	2127
F	528	2186	415	2529

Table 8. Raw mean F1 and F2 frequencies of BBC newsreaders and Serbian speakers

3.2.3 GOAT

Looking at Figure 18 we can see that the BBC newsreaders' GOAT onset is in the centre of the vowel space, overlapping with NURSE. Its glide also ends in a central position, only slightly fronter than the onset.

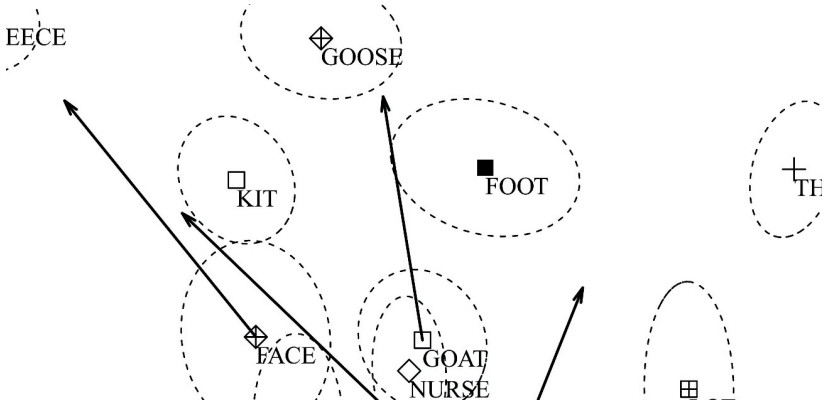


Figure 19. BBC newsreaders' front mid portion of the vowel space; mean formant values were normalised according to Lobanov (1971); ellipses are one standard deviation

On the other hand, as Figure 20 shows, the Serbian speakers' GOAT is retracted, both in terms of the nucleus and in terms of the glide. This can be ascribed either to the influence of North American English, or perhaps to L1 influence combined with the spelling (i.e. identification of the onset with the Serbian /o/), or to both.

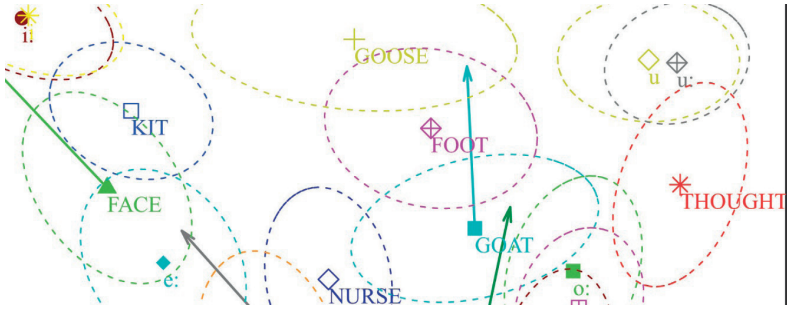


Figure 20. Serbian participants' vowels; mean formant values were normalised according to Lobanov (1971); ellipses are one standard deviation.

Looking at individual charts not shown here we see, similar to what we saw with PRICE/MOUTH, that Serbian participants who display fronter, more native-like GOAT quality are S12, S14, S15, S17, S19 and S21 (these all belong to the older, more experienced group except S12). This is confirmed in Figure 21, where we separate the two groups of Serbian participants.

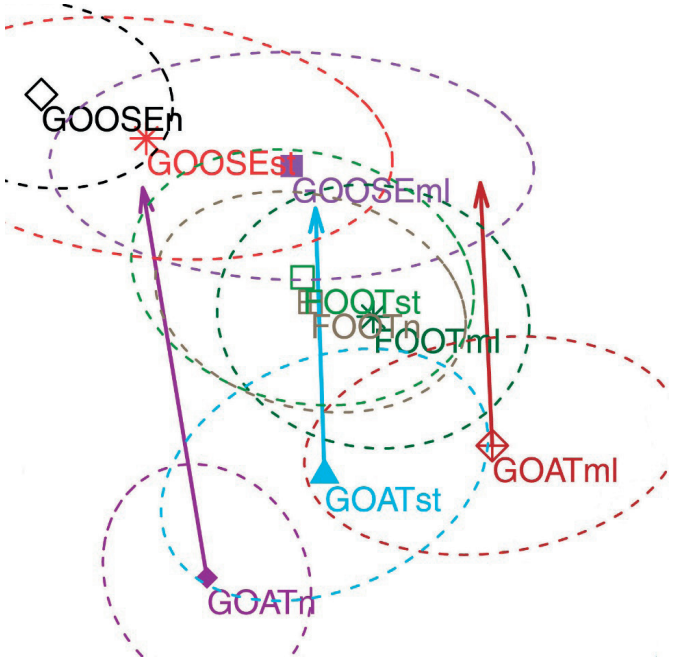


Figure 21. Mean formant values of less experienced (GOATml, FOOTml, GOOSEml) and more experienced (GOATst, FOOTst, GOOSEst) Serbian speakers' English vowels, as well as BBC newsreaders' (GOATn, FOOTn, GOOSEn) vowels, normalised according to Lobanov (1971) (ellipses of other vowels, such as LOT and THOUGHT, have been removed for the sake of clarity).

Figure 21 shows that both in terms of the onset and in terms of the glide the older, more experienced students' GOAT vowel (GOATst) is more like that of the native speakers, and halfway between theirs and the GOAT vowel of the first-year participants.

	GOAT ons.		GOAT glide	
BBCM	493	1437	349	1523
BBCF	568	1685	388	1752
M	487	1200	392	1190
F	555	1380	439	1414

Table 9. Raw mean F1 and F2 frequencies of BBC newsreaders and Serbian speakers

4. STATISTICAL ANALYSIS

Here we show the results of the statistical analysis that aimed to determine the strength of the effect of belonging to the first-year participant group as opposed to the final year student group. Rbrul, v. 3.1.2 (Johnson 2009, Johnson 2017) was used, with the independent variable being whether a speaker was 'ml' (S1–S13) or 'st' (S14–S26), and the dependent variables being normalised F1 and F2 values.

Table 10 shows p values (for $p<0.001$) in ascending order, as well as R^2 values; in other words, vowels that are near the top of the table were significantly different in the speech of two participant groups.

	p	R2
GOAT F2	0.00000000000000000278	0.173
GOATgl F2	0.0000000000000000192	0.156
PRICE F2	0.00000000289	0.064
GOOSE F2	0.00000000314	0.0784
PRICEgl F1	0.0000000185	0.0576
PRICEgl F2	0.0000000531	0.054
MOUTHgl F1	0.0000000807	0.0696
THOUGHT F2	0.000000244	0.0441
MOUTH F2	0.00000167	0.0559
FACE F1	0.0000234	0.0333
MOUTHgl F2	0.0000371	0.0417
THOUGHT F1	0.0000499	0.0275
KIT F1	0.0000671	0.0263
FOOT F2	0.000103	0.0412
STRUT F1	0.000336	0.0295
NURSE F2	0.000359	0.0248
LOT F2	0.000859	0.0205

Table 10. Formants of vowels arranged from the smallest to the largest p value, for $p<0.001$ (the independent variable was group membership).

5. DISCUSSION AND CONCLUSION

Table 10 above shows that where the two groups of student participants differed the most was the following: the F2 of GOAT (both onset and glide), with the more experienced students showing fronter, more native-like values; F2 of GOOSE of which the same is true; F2 of the PRICE and MOUTH onsets, with the older group having a fronter MOUTH onset and backer PRICE onset (again, more native-like and showing a clear separation of the two); and to a somewhat lesser degree F2 of THOUGHT (with older students again showing somewhat closer and more retracted, i.e. more native-like qualities)

On the whole, our results suggest that Serbian-speaking learners, at the proficiency level of our participants, do not on the whole substitute their L2 vowels with L1 vowels; on the other hand, they also do not quite reach the qualities characteristic of native speakers, but rather reach compromise values. Exceptions, regarding substitution, are DRESS, and for some informants at least, TRAP (both are substituted by the Serbian short /e/). The LOT vowel occasionally exhibited American influence, which resulted in a more open quality (more open than both the closest Serbian equivalent and the contemporary British quality); this is perhaps due to the American LOT quality being more readily perceived as a 'new' quality, which according to Flege's Speech Learning Model, is acquired more readily (Flege 1995). Regarding diphthongs, a certain degree of substitution was noticed, with the nucleus of FACE being substituted with the Serbian /e/ by some of the informants, and the nuclei of PRICE and MOUTH being substituted with the Serbian /a/.

As regards Flege and Bohn's hypothesis (Bohn and Flege 1992) that predicts that 'similar' sounds will be acquired more or less equally well by more experienced and less experienced learners, while 'new' sounds will be acquired better by more experienced learners, our results only partially confirm it. Namely, FLEECE and DRESS, being very similar to the Serbian /i:/ and /e/ are indeed much the same for our two groups of participants; on the other hand our more experienced participants acquired the GOOSE vowel better, and [ʉ] is indeed a 'new sound' for Serbian speakers. However, our results regarding TRAP and THOUGHT do not confirm this hypothesis — we would classify the latter as a 'similar' sound (very close to the Serbian /o:/) and yet the more experienced group had a more native-like THOUGHT; on the other hand the TRAP vowel, albeit a 'new' sound, is very similar for our two groups of participants. Finally, our results regarding NURSE are inconclusive, as it is a 'new sound' for Serbian speakers, but the difference between the two groups of participants is fairly slight (more experienced participants' NURSE is somewhat fronter, reaching the native speakers' values, but the F1 difference, characteristic of less experienced speakers, largely remains).

REFERENCES

Bjelaković, A. 2017. The Vowels of Contemporary RP: Vowel Formant Measurements for BBC Newsreaders. *English Language and Linguistics* 21(3), 501–532.

Boersma, P. 2013. Acoustic analysis. In R. J. Podesva & D. Sharma (eds.) *Research methods in linguistics*. Cambridge: Cambridge University Press, 375–397.

Boersma, P. & D. Weenink. 2014. *Praat: Doing phonetics by computer* (version 5.4). [Internet]. Available at: www.praat.org. [14.3.2018].

Bohn, O.-S. & J. E. Flege. 1992. The Production of New and Similar Vowels by Adult German Learners of English. *Studies in Second Language Acquisition* 14, 131–158.

Cruttenden, A. 2014. *Gimson's Pronunciation of English*. 8th edition. London and New York: Routledge.

Čubrović, B. 2016. *Acoustic Investigations of Serbian and American English Vowel Inventories*. Beograd: Filološki fakultet.

Čubrović, B. 2017. Low Back Merger in Native and Nonnative Speakers of American English. *The Linguistics Journal* 11(1), 222–231.

Dančetović, N. & I. Nešić. 2017. Akustička analiza produkcije engleskih monoftonga na tercijarnom nivou. *Nasleđe* 36, 261–283.

Flege, J. E. 1988. The Production and Perception of Foreign Language Speech Sounds. In H. Winitz (ed.) *Human Communication and Its Disorders: A Review*. Norwood, NJ: Ablex, 224–401.

Flege, J. E. 1991. Perception and Production: The Relevance of Phonetic Input to L2 Phonological Learning. In C. Ferguson & T. Huebner (eds.) *Crosscurrents in Second Language Acquisition and Linguistic Theories*. Philadelphia: John Benjamins, 249–289.

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. 2005. Origins and development of the Speech Learning Model. 1st ASA Workshop on L2 Speech Learning, Simon Fraser University, Vancouver, BC, 14–15 April, 2005.

Flege, J. E., O.-S. Bohn & S. Jang. 1997. Effects of Experience on Non-Native Speakers' Production and Perception of English Vowels. *Journal of Phonetics* 25, 437–470.

Jakobson, R. [1937] 1962. On the Identification of Phonemic Entities. *Selected Writings I*. The Hague: Mouton, 418–425.

Johnson, D. E. 2009. Getting Off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3 (1), 359–383.

Johnson, D. E. 2017. Rbrul version 3.1.2. [Internet]. Available at: http://www.danielezrajohnson.com/rbrul.html [14.3.2018]

Krebs-Lazendic, L. 2008. *Early vs. Late Serbian-English Bilinguals' Responses to Two Australian English Vowel Contrasts*. Unpublished PhD dissertation, University of Western Sydney. [Internet]. Available at: http://researchdirect.uws.edu.au/islandora/object/uws%3A5890. [14.3.2018.]

Krebs-Lazendic, L. & C. Best. 2007. Early and Late Bilinguals' Vowel Perception and Production: English Vowel Contrasts that Give Serbian-English Bilinguals a H(E)

AD-ache. In A. S. Rauber, M. A. Watkins & B. O. Baptista (eds.) *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*. 282–292.

Krebs-Lazendic, L. and C. Best. 2013. First Language Suprasegmentally-Conditioned Syllable Length Distinctions Influence Perception and Production of Second Language Vowel Contrasts. *Laboratory Phonology* 4 (2), 435–474.

Labov, W., S. Ash & C. Boberg. 2006. *The Atlas of North American English*. Berlin: Mouton de Gruyter.

Lehiste, I. & P. Ivić. 1986. *Word and Sentence prosody in Serbo-Croatian*. Cambridge, MA: MIT Press.

Lobanov, B. M. 1971. Classification of Russian Vowels Spoken by Different Speakers. *Journal of the Acoustical Society of America* 49, 606–608.

Marković, M. 2007. *Kontrastivna analiza akustičkih i artikulacionih karakteristika vokalskih sistema engleskog i srpskog jezika*. Unpublished doctoral dissertation. Univerzitet u Novom Sadu.

Marković, M. 2009a. Different Strategies in Acquiring L2 Vowels: The Production of High English Vowels /iː, ɪ, uː, ʊ/ by Native Speakers of Serbian. In B. Čubrović & T. Paunović (eds.) *Ta(l)king English Phonetics Across Frontiers*. Cambridge: Cambridge Scholars Publishing, 3–18.

Marković, M. 2009b. The perception and production of the English vowels /e/ and /æ/ by native speakers of Serbian. In A. Tsangalidis (ed.) *Selected papers from the 18th International Symposium of Theoretical and Applied Linguistics*. Thessaloniki: Aristotle University of Thessaloniki, 253–262.

Marković, M. 2012. *Uporedna proučavanja vokala engleskog i srpskog jezika: između univerzalnog i specifičnog*. Novi Sad: Filozofski fakultet.

Marković M. & B. Jakovljević. 2013. The Acquisition of English Diphthongs by Advanced Serbian ESL Learners: A Production Study. *Zbornik Matice srpske za filologiju i lingvistiku* 56 (2), 117–130.

Marković, M. & 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* 59 (2), 97–108.

Mlinar, R. 2011. *Pronunciation of English Diphthongs by Speakers of Serbian: Acoustic Characteristics*. Master's thesis. University of Novi Sad.

Paunović, T. 2002. *Fonetsko-fonološka interferencija srpskog jezika u percepciji i produkciji engleskih vokala*. Unpublished doctoral dissertation. Univerzitet u Nišu.

Paunović, T. 2011. Sounds Serbian? Acoustic Properties of Serbian EFL Students' Speech. In E. Kitis, N. Lavidas, N. Topintzi & T. Tsangalidis (eds.) *Selected Papers from the 19th International Symposium on Theoretical and Applied Linguistics*. Thessaloniki: Aristotle University of Thessaloniki, School of English, Department of Theoretical & Applied Linguistics, 357–369.

Sudimac, N. 2016. Kontrastivna analiza visokih/zatvorenih vokala u produkciji izvornih govornika britanskog engleskog i srpskog jezika. *Filolog* 14, 36–55.

Thomas, E. R. & T. Kendall. 2007. NORM: The vowel normalization and plotting suite. [Internet]. Available at: http://lingtools.uoregon.edu/norm/norm1.php [14.3.2018].

Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Wells, J. C. 1997. Whatever Happened to Received Pronunciation? In C. Medina Casado & C. Soto Palomo (eds.) *II jornadas de estudios ingleses*. Universidad de Jaén, 19–28. http://www.phon.ucl.ac.uk/home/wells/rphappened.htm

Windsor Lewis, J. 1995. Changes in British English Pronunciation During the Twentieth Century. *English Language Teaching News* 25, 88–89. http://www.yek.me.uk/changestwe.html

SUMMARY

FORMANT MEASUREMENTS OF SERBIAN SPEAKERS' ENGLISH VOWELS

We present the results of an acoustic investigation of English vowels as produced by Serbian speakers, students in the English Department, Faculty of Philology, University of Belgrade. The number of participants was 26 (13 first-year students, and 13 fourth-year/MA students), and measured were F1 and F2 of stressed vowels — 11 monophthongs (KIT, DRESS, TRAP, FOOT, STRUT, LOT, FLEECE, PALM, GOOSE, THOUGHT and NURSE) and 4 diphthongs (GOAT, PRICE, MOUTH and FACE). Measurements were also made of the participants' L1 vowels, with which their L2 vowels were compared. Participants were recorded reading BBC news copy in English, and two very short stories in Serbian. The number of tokens analyzed was 7534 for English (around 305 per speaker), and 4266 for Serbian. The results show that Serbian-speaking learners, at the proficiency level of our informants, do not on the whole substitute their L2 vowels with L1 vowels; on the other hand, they also do not quite reach the qualities characteristic of native speakers, but rather reach compromise values. Exceptions, regarding substitution, are DRESS, and for some informants at least, TRAP (both are substituted by the Serbian short /e/). The vowels that exhibited the largest intergroup differences were GOAT, GOOSE, MOUTH, PRICE, and to a lesser degree THOUGHT, with older students showing more native-like qualities.

KEYWORDS: RP vowels, L2 production, vowel quality, EFL students' vowels, formant frequency, Serbian.

ARTICLE INFO:
Original research article
Received: June 15, 2018
Revised: February 1, 2019
Accepted: February 13, 2019

UDC: 811.111'342.9
811.14'342.9
DOI: https://doi.org/10.18485/philologia.2018.16.16.3

L2 ACQUISITION AND PRODUCTION OF THE ENGLISH RHOTIC BY L1 GREEK-CYPRIOT SPEAKERS: THE EFFECT OF L1 ARTICULATORY ROUTINES AND PHONETIC CONTEXT

DIMITRA DIMITRIOU[1]
University of Central Lancashire
UK

U ovoj studiji se ispituje produkcija engleskog rotičkog glasa kod govornika kiparskog grčkog, čija je tipična realizacija u L1 alveolarni dotačnik (eng. *tap*). Takođe se porede produkcija učenika koji pohađaju školu na engleskom i onih koji pohađaju školu na grčkom tokom srednjoškolskog obrazovanja. Ispitanici su studenti koji su bili uzrasta 17–21 kada su se preselili u Veliku Britaniju; dužina boravka bila je 1–4 godine, a starost ispitanika 18–24 godina. Kontrolnu grupu je sačinjavalo šestoro izvornih govornika engleskog. Učesnici su snimljeni dok izgovaraju liste reči na grčkom i engleskom jeziku sa rotičnim glasom u početnoj, međuvokalskoj poziciji i u grupama Cr i rC. Ispitivani su način tvorbe, trajanje i frekvencija formanata. Rezultati ukazuju na to da su učenici iz škola na engleskom jeziku uspešniji u produkciji engleskog aproksimanta nego polaznici škola na grčkom, premda nijedna od dve grupe ispitanika ne doseže vrednosti izvornih govornika u svim kontekstima, ni kada je reč o trajanju, ni kada je reč o frekvenciji formanata. Takođe su primećeni efekti fonetskog sistema L1 na produkciju rotičnog glasa u L2. Ova studija daje uvid u oblast kojoj je pruženo malo pažnje u kontekstu Kipra, i predstavlja osnov za buduće istraživanje koje može dovesti do poboljšanja u predavanju i učenju engleskog na Kipru i drugim jezicima sa sličnim fonetskim inventarom.

Ključne reči: L2 produkcija rotičkog glasa, SLM, asimilacija fonetske kategorije, disimilacija fonetske kategorije, interakcija L1-L2, grčko-kiparski učenici engleskog.

1 Contact information: dimitriou.dimitra18@gmail.com

1. INTRODUCTION

Previous studies on the acquisition and production of the rhotic in a second language (L2) by groups of learners from various first-language (L1) backgrounds (e.g. Chan 2010; Rose 2010; Olsen 2012) demonstrate that rhotics are problematic in L2 speech acquisition, especially for learners whose L1 and target L2 involve different realisations. However, to my knowledge, no previous study investigated the production of the rhotic approximant by native Greek or Greek-Cypriot learners of English, whose L1 involves a tap realisation of the rhotic. Therefore, the purpose of this study was to examine the acquisition of the rhotic approximant by native Greek-Cypriot learners of English, in order to identify the difficulties that these learners face in its production.

The task of the learner in acquiring the sound patterns of an L2 in perception and production is complex and subject to various constraints. The question of how the two subsystems of a bilingual interact is addressed by Flege (1995) in his Speech Learning Model (SLM), according to which the L1 and L2 phonetic categories mutually influence one another through the processes of phonetic category assimilation and phonetic category dissimilation. Phonetic category assimilation occurs when the establishment of a new category is blocked due to the perception of an L2 sound as phonetically similar to an L1 sound, at least in the early stages of learning (Flege 1995). Phonetic category dissimilation occurs when a new phonetic category is ultimately formed for an L2 sound (Flege 1995). By hypothesis, whether a new category will be formed for an L2 sound depends on the degree of development of a neighbouring L1 sound, and the perceived phonetic dissimilarity of an L2 sound with the closest L1 sound (Flege 2007).

In addition, the SLM hypothesises that perception is linked to production in a way that difficulties in perception lead to difficulties in production, even though segmental production and perception are not necessarily brought into perfect alignment, as motor programs are also needed for successful production (Flege 2003). This means that learners may perceive the phonological characteristics of an L2 sound, but still have an inadequate knowledge of the motor commands required for its articulation (Leather and James 1991).

2. THE RHOTIC IN STANDARD MODERN GREEK (SMG) AND CYPRIOT-GREEK (CYG)[2]

The SMG rhotic is typically realised as an alveolar tap but can also occur as a short trill with two or three cycles (Arvaniti 1999a). Baltazani and Nicolaidis (2012; 2013) and Nicolaidis and Baltazani (2011; 2014) examined the production of the rhotic in SMG, and found that it consists of two components: a single constriction and a vowel-like transition (vocoid). The position of the vocoid was found to vary depending on its position; it appears between the constriction and the preceding or following consonant in Cr (Consonant-r) and rC (r-Consonant) clusters, and before the single constriction in

2 The terms "Greek-Cypriot" and "Cypriot-Greek" can be used interchangeably, although it is more common for the former to refer to people and the latter to refer to the dialect.

phrase-initial position (Baltazani and Nicolaidis 2013). The duration of the rhotic was found to be on average 55-60 ms in consonant clusters (both components) (Baltazani 2009) and 23.3 ms in intervocalic position (only the constriction phase). In phrase-initial position, the constriction was found to be on average 24.7 ms long, while the vocoid was more than double in length (50.43 ms) in this position (Baltazani and Nicolaidis 2012).

The degree of constriction of the rhotic was found to vary from complete to incomplete contact and very open articulations. Overall, incomplete constrictions were found in 47% of the tokens in tautosyllabic Cr clusters, 57% in heterosyllabic rC sequences, and 63% in initial and intervocalic positions (Baltazani and Nicolaidis 2013). In addition, in rC and Cr clusters, more tokens were produced with incomplete constriction in the context of a fricative compared to a stop (Fricative-r: 49% and r-Fricative: 67% compared to Stop-r: 44% and r-Stop 47%). Finally, Baltazani (2005) briefly mentions the presence of frication noise in some tokens resulting from intermediate degrees of constriction. As regards the vocoid, the overwhelming majority of tokens (80%) were produced with a modal or breathy voice, especially in Cr and phrase-initial contexts (Nicolaidis and Baltazani 2015).

Similarly to SMG, the rhotic in CyG is typically produced as a tap when single; however, Arvaniti (1999b; 2001a) and Arvaniti and Tserdanelis (2000) argue that the CyG system also contains a trill articulation of the rhotic when geminated. This is different from the realisations in SMG, which has no geminate productions. However, in the comparison between single segments in SMG and CyG, Arvaniti (2001b) found that the single /ɾ/ was similar in duration in both varieties across speaking rates, which was expected since the tap is not free to shorten or lengthen at will. Overall, the realisation of the rhotic was found to be very similar in both varieties. Additionally, the lexical items used in this study are found in the vocabulary of both varieties and the Greek wordlist was written according to the SMG spelling and grammatical rules as CyG has no established orthography. Therefore, the results of this study may be at least to some extent generalised to SMG speakers as well.

3. PRESENT STUDY

This study focused on the production of the rhotic approximant by native Greek-Cypriot learners of English, in order to assess whether these learners face problems in its acquisition. In addition to investigating the acquisition of the rhotic approximant by learners whose L1 involves a tap, this study also compared two groups of Greek-Cypriot learners of English, namely learners that had attended English-speaking and learners that had attended Greek-speaking schools during their secondary education. More specifically, this study addressed the following research questions:

1. Does the phonetic system of the Greek-Cypriot dialect affect the production of the rhotic in English, and if so, in what ways?
    a. In what contexts is the rhotic produced more accurately by Greek-Cypriots in terms of manner of articulation, and what are the possible reasons for that?

   b. How do Greek-Cypriot learners of English who study in the UK produce the rhotic approximant in different contexts, and how are their productions different from the productions of native English speakers in terms of duration and tongue configurations?
2. Is there a difference in the production of the rhotic between Greek-Cypriots that attended English-speaking private schools, and those that attended Greek-speaking public schools, and if so, why?

4. METHODOLOGY AND PROCEDURES

4.1 SPEAKERS

In order to address the research questions of this study, three groups of participants were required. Group A consisted of 13 Greek-Cypriot speakers who had attended public Greek-speaking schools in Cyprus during their primary and secondary education. The 9 participants in Group B had attended private English-speaking schools during their secondary education. All Greek-Cypriot participants had started learning English as a Foreign Language at an average age of 8 years. They had all obtained an IGCSE or IELTS certificate in order to be admitted to a UK university, and were therefore considered to be competent users of the English language. Participants in Groups A and B had arrived in the UK at an average age of 18.5 and had lived there for an average of 2.4 years. Their age ranged from 18-24 years old (average 20.6). The two CyG groups differed in their language use patterns during secondary school, as obtained through self-reports on a seven-point scale, but reported similar language use patterns during their university years. Table 1 presents the average values and range of values obtained by Greek-Cypriot participants regarding their language use patterns. Group C was a control group and consisted of 6 monolingual native speakers of English (age range 19-22) who had been born, brought up and lived in Lancashire, UK at the time, chosen so as to avoid regional variation in the production of the rhotic. All participants were students at Lancaster University. More details concerning participants' characteristics can be found in Table A1 in the Appendix.

		Group A		Group B	
		Average	Range	Average	Range
L1 USE	**School: class**	6.85	6-7	3.5	2-5
	School: social	6.92	6-7	5.75	5-7
	Uni: class	2.23	1-5	2.63	1-5
	Uni: social	5.08	4-6	5.5	5-7
L2 USE (School)	**Class**	2.31	1-3	5.88	5-7
	Social	1.38	1-3	2.75	1-5
	Class (w. native)	1.15	1-3	3.75	1-6
	Social (w. native)	1.15	1-3	3	1-5
	Class (w. non-native)	1.46	1-3	2.5	1-7
	Social (w. non-native)	1.23	1-3	2.13	1-7
L2 USE (University)	**Class**	5.23	4-6	5.63	5-7
	Social	4.15	3-6	3.63	2-6
	Class (w. native)	4.15	1-7	3.88	2-7
	Social (w. native)	3.08	1-5	3.5	2-7
	Class (w. non-native)	3.69	1-7	3.88	1-7
	Social (w. non-native)	3.62	1-6	3.75	1-6
Motivation	**Importance of pronunciation**	5.85	3-7	5.63	4-7
	Attention paid	4.77	3-7	5.75	4-7

Table 1. Language use patterns of Greek-Cypriot participants based on self-reports

4.2 SPEECH MATERIAL AND RECORDING SESSIONS

Participants in Groups A and B were required to produce 16 Greek and 21 English tokens included in the carrier phrase "_____ mu ipe ki efige" ("_____ he said and left") and interspersed with fillers. Group C only read the English phrases. The tokens recorded contain a rhotic sound in six different contexts (Table 2). The same consonant clusters were used for both languages, apart from clusters that do not have an equivalent real Greek word. The same tautosyllabic vowel or variant was used for English tokens and

their Greek equivalent[3]. Each phrase was presented separately using PowerPoint, and the process was repeated twice. Participant B8 was excluded from the study due to poor production caused by illness, rendering their productions inaudible. The total number of tokens analysed was 1134 English words (27 participants x 21 tokens x 2 repetitions) and 672 Greek words (21 participants x 16 tokens x 2 repetitions).

Recordings were carried out in comfortable and quiet environments at Lancaster University, individually or in groups of maximum 3 participants. The recorder used was a MicroTrack II 2-Channel Digital Recorder. The sessions lasted approximately 30 minutes for Greek-Cypriot speakers and 15 minutes for native English speakers.

Context	English	Greek
(1) Word-Initial	Rich	Rito ([ɾiˈtɔ] 'a saying')
(2) Word-Medial/ Intervocalic	Very	Mesimeri ([mɛsiˈmɛɾi] 'noon')
(3a) Cr clusters with voiced and (3b) voiceless stops	Brick Drink Grim	Brizola ([bɾiˈzɔlɐ] 'steak') Dripla ([ˈdɾiplɐ] 'dribbling') Grimatsa ([gɾiˈmɐtsɐ] 'grimace')
	Priest Tree Cream	Prin ([ˈpɾin] 'before') Triti ([ˈtɾiti] 'Tuesday) Krima ([ˈkɾimɐ] 'shame')
(4) Cr clusters with voiceless fricative	Free Three **Shrink**	Friki ([ˈfɾiki] 'horror') Thrilos ([ˈθɾilɔs] 'legend') --
(5a) rC clusters with voiced and (5b) voiceless stops	**Orbit** **Ordination** **Organised**	-- -- --
	Harpoon Artistic Arcade	Arpa ([ˈɐrpɐ] 'harp') Artios ([ˈɐrtiɔs] 'even (number)') Arketa ([ɐrkɛˈtɐ] 'enough')
(6) rC clusters with voiceless fricatives	Surfing Earth Arson **Harsh**	Aderfi ([ɐðɛɾˈfi] 'sister' Ipertheama ([ipɛɾˈθɛɐmɐ] 'spectacle') Arseniko ([ɐrsɛniˈkɔ] 'male' or 'arsenic') --

Table 2. Tokens recorded in Greek and English, grouped according to the context of the rhotic

3 SMG and CyG have a 5-vowel system: /i, ɛ, ɐ, ɔ, u/. For a description of the quality of SMG vowels, see Arvaniti (1999a; 2007) and for CyG, Arvaniti (1999b).

4.3 MEASUREMENTS

The analysis was carried out using Praat version 5.3.42. Rhotics were firstly classified according to their manner of articulation (approximant, tap, trill) based on auditory and acoustic analyses. One additional variant of the tap was discovered during analysis and was labelled as a tap followed by frication noise. Formant values at 50% were also taken. To measure its duration, the onset and offset of the rhotic were determined by changes on the spectrogram and/or waveform suggesting the onset or offset of the surrounding segments, combined with auditory analysis. The onset of word-initial taps and taps in Cr clusters was determined at the beginning of the vocoid. Care was taken to ensure that measurements were consistent across participants and tokens. An intra-rater reliability assessment of the measurements was carried out on a 10% representative sample of the data. The average absolute difference ranged from 2-8 ms and was considered acceptable; therefore, the original set of measurements was used for analysis.

Clusters with voiceless stops preceding the rhotic were excluded from duration and formant frequencies analyses due to the fact that as opposed to Greek /p, t, k/, English voiceless stops are produced with aspiration in syllable-initial position and this had different effects for Greek-Cypriot learners compared to native speakers. Duration and formant values for rC clusters are also not reported due to the lack of native speaker productions with which to make comparisons.

4.4 HYPOTHESIS

It was hypothesised that the highest proportion of approximants would be produced in Fricative-r and r-Fricative clusters in English by Greek-Cypriot speakers, followed by clusters with voiced and voiceless stops, and finally, by the rhotic in initial and intervocalic position. This hypothesis was based on the degree of constriction of the rhotic in various contexts in SMG, as examined by Baltazani and Nicolaidis (2013) (see section 2). In addition, even though in initial and intervocalic position the rhotic was found to be regularly produced with incomplete constriction in Baltazani and Nicolaidis (2013), Cruttenden and Gimson (2014: 227) argue that, "the approximant in initial position may be the most troublesome articulation of all" for foreign language learners. Therefore, I expected Greek-Cypriot learners to have some difficulty in producing the approximant in this position. Finally, where there is no equivalent CyG cluster, it was expected that the pronunciation would be more native-like. "Harsh" and "Shrink" were expected to be more accurately pronounced since the position of the tongue in /ʃ/ approximates more the position for the production of the English rhotic.

5. RESULTS

5.1 MANNER OF ARTICULATION

One interesting result of the analysis was the production of the tap with frication noise in both Greek and English words, which has not been extensively discussed in the literature. These were instances where the constriction phase of the tap was followed by frication noise instead of the expected vocoid phase. Taps with frication noise only occurred in r-Fricative and r-(Voiceless)Stop clusters in Greek words and were the most common variant in these contexts by both Groups A and B. In English, the occurrence of this variant was reduced, especially for Group B; however, its occurrence in English was more widespread across contexts, with the most instances of this variant found in (Voiceless)Stop-r and r-(Voiceless)Stop clusters for both groups.

Tables 3 and 4 below show the percentage of occurrence of each variant in all contexts in English and Greek respectively as produced by Greek-Cypriot speakers. A quick overview reveals that both groups used more approximants in total than any other variant, with Group B producing more approximants than Group A overall. However, in combination with the taps with frication noise, the tap was used more often than approximants by participants in Group A making up for more than 53% of the productions for this group. Group B produced considerably more approximants than any other variant. This was also the case in rC clusters, where elision takes place for Group C. Group A on the other hand demonstrated greater variation. Interestingly, elision in rC clusters occurred slightly more often in the productions of Group A than Group B. However, approximant productions in these contexts were considered as accurate during the analysis, as elision is dependent on the variety of English used, despite the fact that the rhotic is omitted in the variety spoken by the control group in this study.

	Approximants		**Taps**		**Taps (Frication)**		**Ø**		Trills	
	A	*B*	*A*	*B*	*A*	*B*	*A*	*B*	*A*	*B*
Word Initial	10/26 (38.46%)	12/16 (75%)	16/26 (61.54%)	4/16 (25%)	-	-	-	-	-	-
Intervocalic	5/26 (19.23%)	10/16 (62.5%)	21/26 (80.77%)	6/16 (37.5%)	-	-	-	-	-	-
Fricative-r	38/78 (48.72%)	38/48 (79.17%)	35/78 (44.87%)	10/48 (20.83%)	4/78 (5.13%)	-	-	-	1/78 (1.28%)	-
r-Fricative	52/104 (50%)	47/64 (73.44%)	19/104 (18.27%)	8/64 (12.5%)	14/104 (13.46%)	4/64 (6.25%)	19/104 (18.27%)	5/64 (7.81%)	-	-
(Voiced)Stop-r	34/78 (43.59%)	36/48 (75%)	43/78 (55.13%)	12/48 (25%)	1/78 (1.28%)	-	-	-	-	-
r-(Voiced)Stop	37/78 (47.44%)	30/48 (62.5%)	24/78 (30.77%)	13/48 (27.08%)	10/78 (12.82%)	1/48 (2.08%)	7/78 (8.97%)	3/48 (6.25%)	-	1/48 (2.08%)
(Voiceless) Stop-r	22/78 (28.21%)	23/48 (47.92%)	23/78 (29.49%)	14/48 (29.17%)	33/78 (42.31%)	11/48 (22.92%)	-	-	-	-
r-(Voiceless) Stop	30/78 (38.46%)	23/48 (47.92%)	14/78 (17.95%)	9/48 (18.75%)	26/78 (33.33%)	9/48 (18.75%)	7/78 (8.97%)	7/48 (14.58%)	1/78 (1.28%)	-
Total	**221/546**	**219/336**	**195/546**	**76/336**	**95/546**	**25/336**	**33/546**	**15/336**	**2/546**	**1/336**
%	**40.48%**	**65.18%**	**35.71%**	**22.62%**	**17.4%**	**7.44%**	**6.04%**	**4.46%**	**0.37%**	**0.3%**

Table 3. Percentage of occurrence of each variant by Greek-Cypriot speakers in English tokens

	Taps		Taps (Frication)		Ø		Trills	
	A	B	A	B	A	B	A	B
Initial	24/26 (92.31%)	16/16 (100%)	-	-	-	-	2/26 (7.69%)	-
Intervocalic	26/26 (100%)	16/16 (100%)	-	-	-	-	-	-
Fricative-r	52/52 (100%)	32/32 (100%)	-	-	-	-	-	-
r-Fricative	15/78 (19.23%)	22/48 (45.83%)	62/78 (79.49%)	25/48 (52.08%)	1/78 (1.28%)	-	-	1/48 (2.08%)
(Voiced) Stop-r	78/78 (100%)	48/48 (100%)	-	-	-	-	-	-
r-(Voiced) Stop	-	-	-	-	-	-	-	-
(Voiceless)Stop-r	77/78 (98.72%)	48/48 (100%)	-	-	-	-	1/78 (1.28%)	-
r-(Voiceless)Stop	25/78 (32.05%)	18/48 (37.5%)	50/78 (64.10%)	25/48 (52.08%)	-	-	3/78 (3.85%)	5/48 (10.42%)
Total	**297/416**	**200/256**	**112/416**	**50/256**	**1/416**	**-**	**6/416**	**6/256**
%	**71.39%**	**78.13%**	**26.92%**	**19.53%**	**0.24%**	**-**	**1.44%**	**2.34%**

Table 4. Percentage of occurrence of each variant by Greek-Cypriot speakers in Greek tokens

As regards the clusters not found in Greek phonology, "Shrink" was the most successful token in terms of manner of articulation for Group A, with the most instances of approximant productions compared to all other tokens across contexts (20/26). "Harsh" was also moderately successfully pronounced, having the most instances of elision in both Group A and B (11/26 and 4/16 respectively), compared to other tokens that native speakers produced with an omission of the /ɹ/. Finally, r-(Voiced)Stop clusters were the second most accurately pronounced cluster for Group A, but fifth for Group B (combining both approximant productions and elision), indicating a higher success rate compared to other contexts for Group A. Approximants (plus elision where appropriate) for the two groups occur most often in the following order:

Group A

1. r-Fricative (68.27%)
2. r-(Voiced)Stop (56.41%)
3. Fricative-r (48.72%)
4. r-(Voiceless)Stop (47.43%)
5. (Voiced)Stop-r (43.59%)
6. Word-Initial (38.46%)
7. (Voiceless)Stop-r (28.21%)
8. Intervocalic (19.23%)

Group B

1. r-Fricative (81.25%)
2. Fricative-r (79.17%)
3. (Voiced)Stop-r=Word-Initial (75%)
4. r-(Voiced)Stop (68.75%)
5. r-(Voiceless)Stop=Intervocalic (62.5%)
6. (Voiceless)Stop-r (47.92%)

5.2 DURATION

Unpaired t-tests were used to test statistical significance of the duration differences among groups in each context. In Fricative-r clusters, both Groups A and B had significant differences with Group C ($p<0.0001$ for Group A and $p=0.012$ for Group B). Duration differences between Group A and Group B were also statistically significant in this case ($p=0.01$). Interestingly, in the individual token "Shrink", Group A had no significant differences compared to Group C ($p=0.196$) whereas Group B did ($p=0.028$). A significant difference was also found between the two Greek-Cypriot groups in this word ($p=0.003$). In general, "Shrink" was more accurately produced by both Group A and Group B in terms of duration in comparison with "Free" and "Three" when the tokens were compared individually. In Stop-r clusters only Group A had significant duration differences ($p=0.033$) with Group C. Groups A and B had no significant differences in these clusters. Group A's approximant duration was generally shorter than the other groups in both Fricative-r and Stop-r clusters. Finally, approximant duration in word-initial and intervocalic positions was also examined, despite the limited number of approximants in these positions. All comparisons between the groups showed no statistical significance in their duration differences.

5.3 FORMANT VALUES

Formant values in both Fricative-r and Stop-r clusters follow a similar pattern, with the F1 slightly higher by Group C as opposed to Groups A and B, and F2 and F3 lower for this group. Again, Groups A and B shared similar ranges and averages, with a very slight tendency for Group B to produce lower F2 and F3 values than Group A. Unpaired t-tests were used to determine whether F1, F2 and F3 differences between the groups were significant in Fricative-r clusters. F1, F2 and F3 differences between Group A and Group B were not statistically significant. Differences in F1, F2 and F3 between Group A and C and Group B and C were found to be extremely statistically significant ($p<0.0001$) in all comparisons. The same pattern was observed for Stop-r clusters, where Groups A and B

had no significant differences whereas significant differences existed between Group A and C and Group B and C in F1, F2 and F3 ($p<0.0001$ in all comparisons).

Finally, in tokens with singleton /ɹ/, F1 differences among the groups were not statistically significant, whereas F2 differences between Group A and C and Group B and C were statistically significant in both "Rich" and "Very" ($p=0.0004$ for Group A and C, $p=0.0016$ for B and C in "Rich"; $p=0.0002$ for Group A and C, $p<0.0001$ for Group B and C in "Very"), but not between Groups A and B. F3 differences were not significant among the three groups in "Rich", but they were in "Very", in which even Groups A and B had significant differences ($p<0.0001$ for A and C, $p=0.0002$ for B and C, $p=0.025$ for A and B).

6. DISCUSSION

6.1 MANNER OF ARTICULATION

The prediction that clusters with fricatives would favour approximant production in English by Greek-Cypriot speakers especially in rC clusters was confirmed for both groups. As mentioned in section 5.1 above, approximant productions in rC clusters were considered as accurate during the analysis of the results, despite the fact that the rhotic is omitted in the variety spoken by the control group in this study. The increased number of approximants found in clusters with fricatives was predicted based on Baltazani and Nicolaidis (2013), who found that in Greek rC and Cr clusters, more tokens were produced with incomplete constriction of the tap in the context of a fricative compared to a stop (see section 2). The next most successfully produced categories for Group A are r-Stop categories, with both voiced and voiceless stops (including both approximant productions and omissions). This was also expected as Baltazani and Nicolaidis (2013) found that rC sequences favour incomplete constrictions of the tap (57% of their tokens produced with incomplete constriction of the tap, compared to 47% in Cr clusters). As predicted, word initial and intervocalic contexts were among the least successful contexts, especially for Group A, along with (Voiceless)Stop-r clusters, which were found to cause difficulties to both groups, perhaps due to the aspiration of word-initial voiceless stops in English.

Group B produced a high percentage of approximants in word-initial position, contrary to the predictions, but not in intervocalic position, which is found at the bottom of the list. It is also important to note that for Group B, (Voiced)Stop-r is the only Cr context that is more successful than its rC equivalent, perhaps because r-(Voiced)Stop clusters are not found in SMG or CyG. In general, with the exception of (Voiceless)Stop-r clusters, the rhotic was produced as an approximant in more than 60% of the tokens in each context for this group, indicating a higher percentage of success compared to Group A. Overall, the results were contrary to the predictions for Group B, probably because this group had more overall experience in English. Their general success rate in all contexts was higher, indicating a process towards achieving complete overall success in the acquisition of the approximant, at least in terms of manner of articulation.

As regards the tokens labelled as taps with frication noise, Baltazani (2005) briefly mentions the occurrence of tokens produced with an intermediate degree of constriction

resulting in frication. In addition, Nicolaidis and Baltazani (2015) observed that while in the majority of their tokens the vocoid phase of the rhotic was produced with a modal or breathy voice (see section 2), there was a large increase of vocoids with whispered quality in rC contexts (over 40% of the tokens in rC clusters), which was interpreted as an assimilatory effect to the following voiceless consonant. Therefore, the frication noise found in Greek rC clusters in this study can be interpreted as the vocoid phase of the tap produced with whispered quality, which extends to the production of the approximant in English rC clusters as well. Its frequency in (Voiceless)Stop-r clusters in English but not Greek tokens may be caused by the fact that in English, syllable-initial voiceless stops are aspirated, as opposed to syllable-initial voiceless stops in Greek. However, the effect of aspiration on rhotic production by English learners whose L1 has unaspirated syllable-initial voiceless stops needs to be further investigated.

Another interesting observation is that while Group B produced significantly more approximants than Group A in all contexts and overall, they did so in rC clusters as well, where the approximant is typically omitted in the productions of native speakers. Interestingly, Group A had a slightly higher percentage of elision in these contexts, especially in the individual token "Harsh". This phenomenon may be the result of input. Having more exposure to English from an earlier age, Group B may have been able to form a new category for the English approximant prior to their arrival in the UK; however, the input may not have been accurate enough to lead to the perception of elision, as teachers in Cyprus may have pronounced the rhotic in such contexts. On the other hand, category formation may not have been achieved by Group A prior to their arrival in the UK, where the native-speaker input may have led to a more British-like perception of elision in these contexts. At the same time, the effect of input by different varieties of English, such as American English may have influenced their representations, since Greek-Cypriot learners are more likely to have received more American-accented input prior to their arrival in the UK through American television programmes that are widely broadcast in Cyprus. In this case, even if both groups had the same amount of American-accented input, Group B participants may have been more likely to form new categories prior to their arrival to the UK, due to the fact that they had more exposure to English earlier in life than Group A participants. Orthographic effects may also affect Greek-Cypriots in general in these clusters, since Greek is a language with a letter-sound correspondence. Further research focusing on the investigation of this observation could provide more insights explaining the effect of input on category formation for this group of learners.

6.2 DURATION

With regard to duration, whereas neither Group A nor Group B achieved overall native-like duration of the approximant, Group B produced longer durations than Group A, approximating more the native productions. Groups A and B only differed significantly in Fricative-r clusters, which suggests that Group B may have achieved a slight modification of the duration values in these clusters. However, the significant differences between Group B and Group C in these clusters indicate that native-like duration was still not achieved by these learners. Group B did, however, achieve native-like approximant duration in Stop-r clusters, where there were no significant

differences compared to Group C. Group A on the other hand was not so successful, since there were significant duration differences in both Fricative-r and Stop-r clusters compared to Group C.

In word-initial and intervocalic positions, neither Groups A nor B had significant differences compared to Group C. The limited number of approximant productions in these contexts suggests that these learners, especially Group A, face difficulties in producing the approximant in these positions, but as soon as they overcome manner of articulation difficulties, they are relatively successful in terms of duration. The generally shorter durations of the two Greek-Cypriot groups, and especially of Group A, may be the result of an effect of the L1 phonetic system on the L2 rhotic acquisition, as the L1 tap is shorter in duration, possibly affecting the realisation of the L2 approximant when it occurs.

6.3 FORMANT VALUES

Formant values in Fricative-r and Stop-r clusters indicate that Greek-Cypriot speakers produce /ɹ/ with a higher and more fronted tongue position, and with less lip rounding than native English speakers. Whether Greek-Cypriots' productions were affected to a greater degree than native speakers' productions by the vowel in the environment of the rhotic (a variant of /i/ in the majority of the tokens) needs to be further explored in future research, since different neighbouring vowels may exert a different influence on the production of the approximant by learners of English.

Overall, formant values indicate that tongue height was more similar between Groups A and B and Group C in Initial and Intervocalic positions than in Fricative-r and Stop-r clusters, but the approximant was still produced as more fronted and with less lip rounding in both words. More native-like productions seem to be achieved in word-initial position followed by intervocalic position, again providing support for the hypothesis that as long as the approximant is acquired, it is produced more accurately in Initial and Intervocalic position than in Cr clusters. However, the data collected for these contexts, as well as the limited number of approximant productions by the Greek-Cypriot participants in these contexts do not allow for any claims to be made with confidence.

6.4 L1-L2 INTERACTION – L1 EFFECTS

Greek-Cypriot speakers were less successful in /ɹ/ production than the Cantonese ESL learners in Chan (2010) (46.52% of approximants plus elision for Group A and 69.64% for Group B, compared with 87.5% for Cantonese learners in Chan 2010). Similarly to Cantonese learners in this study, and contrary to the SLM hypothesis, the absence of an approximant in the CyG and SMG phonological system does not seem to facilitate its acquisition and production in the L2. Olsen (2012) reports the same phenomenon for L1 English learners of Spanish, who produce the Spanish rhotic as an approximant instead of a tap. This is the result of the L2 category being affected "by the phonological

structure of those L1 categories to which they are most similar, at least in the early stages of L2 development" (Olsen 2012: 70). In the case of Greek-Cypriot learners, and based on the assumptions of the SLM, approximants may at the first stages of learning be perceived as the CyG tap, blocking the establishment of a new category, due to a small but sufficient perceived phonetic similarity between them.

However, there is no reason to assume that the CyG tap and the English approximant are perceived to be the same sound, at least after the initial stages of learning, as the two sounds are inherently different. In theory, based on the fact that taps and approximants are very different, and the phonetic inventory of (Cypriot-)Greek does not contain an approximant rhotic, native (Cypriot-)Greek learners should be able to form a new category for the English rhotic (see section 1). However, this does not mean that they will produce the approximant instead of the tap in their speech, neither that they will do so accurately, due to articulatory difficulties, or due to the added effort of producing an approximant instead of the well-known articulatory configurations for a tap (Leather and James 1991; Flege 2003). When the approximant is used, its production is different than the native productions in terms of duration and tongue position, providing support for the argument that while learners may perceive the phonological characteristics of an L2 sound, they may still face motor constraints in its articulation. Seeing that the participants in Group B, who had more experience in English and had used it more often in earlier stages compared to Group A participants, produced higher rhotic accuracy rates, there is no reason to assume that Group A participants will not eventually achieve the same accuracy.

The success in the production of clusters that have no Greek equivalent (i.e. Cr and rC clusters with /ʃ/ and rC clusters with voiced stops) especially by Group A, points to an explanation of new category formation. Based on the assumptions of the SLM, since no Greek equivalent exists to cause transfer, a new category is more likely to be formed for the rhotic in these clusters, which will reflect the native input that the learners had experienced. In addition, the consonantal environment of the rhotic in these tokens favoured approximant production due to the similarity in place of articulation between /ʃ/ and /ɹ/. At the same time, their apparently limited success may be due to the learners having received inadequate amounts of input due to their L1-L2 use patterns during their stay in the UK. With regard to r-(Voiced)Stop clusters, although they seem to be moderately successful in terms of manner of articulation especially for Group A, the extensive use of taps and taps with frication noise (43.59% for Group A and 29.16% for Group B) calls for further investigation of the success rate in this context.

7. CONCLUSION

This study investigated rhotic production by native Greek-Cypriot speakers who had started learning English as an L2 at an early age but had arrived in a predominantly English-speaking country late in life, in early adulthood. The research questions investigated were the following:

1. Does the phonetic system of the Greek-Cypriot dialect affect the production of the rhotic in English, and if so, in what ways?

a. In what contexts is the rhotic produced more accurately by Greek-Cypriots in terms of manner of articulation, and what are the possible reasons for that?
b. How do Greek-Cypriot learners of English who study in the UK produce the rhotic approximant in different contexts, and how are their productions different from the productions of native English speakers in terms of duration and tongue configurations?

2. Is there a difference in the production of the rhotic between Greek-Cypriots that attended English-speaking private schools, and those that attended Greek-speaking public schools, and if so, why?

To summarise the findings concerning research question 1a, differences were found between the two Greek-Cypriot groups as to the contexts in which the approximant was more accurately produced, with a general tendency for clusters with fricatives to favour approximant production. Word-initial /ɹ/ had a different success rate for the two Greek-Cypriot groups, whereas intervocalic /ɹ/ was the least successful for Group A and second to last for Group B. Clusters with voiceless stops before the rhotic were among the least successful by both groups as a possible result of the aspiration differences in syllable-initial voiceless stops between the two languages, and were therefore excluded from duration and formant frequencies analyses. The relative success of some but not other contexts was attributed to the degree of constriction of the tap in the equivalent contexts in CyG and SMG on the one hand, and to ease of articulation on the other.

With regard to research question 1b, neither Group A nor Group B achieved overall native-like duration of the approximant in all contexts. Group B showed a tendency to approximate native-like durations especially in (Voiced)Stop-r clusters. Group A had significant differences compared to the native speakers in both Fricative-r and (Voiced)Stop-r clusters. It was also found that in word-initial, intervocalic and partially in (Voiced)Stop-r positions, Greek-Cypriot speakers are more successful in achieving native-like duration, but only after they overcome manner of articulation difficulties, as opposed to clusters with fricatives in which approximants are generally more frequent but native-like duration is not achieved. In terms of formant frequencies, Greek-Cypriots showed a possible influence of the tautosyllabic vowel on rhotic production, with their approximants produced higher, more fronted and with less lip rounding than those of the native speakers. Again, "Rich" and "Very" seem to be more successful than Fricative-r and Stop-r clusters in terms of tongue configuration, providing support for the hypothesis that as long as the approximant is acquired in these contexts, more accurate duration and formant values are achieved.

Concerning research question 2, apart from the differences already discussed, students that attended English-speaking schools (Group B) appear to perceive and produce the English approximant more accurately than the students that attended Greek-speaking schools (Group A). In general, Group B produces considerably more approximants than taps in all contexts and more than Group A. However, whereas students from an English-speaking school background are more successful in terms of the duration and tongue configurations for the approximant, both groups have yet to

acquire the specific phonetic features of the English approximant as produced by native speakers of English. Group B's relative success compared to Group A was attributed to their higher degree of exposure to English as well as more native-speaker input during the early stages of learning, which may have resulted in new category formation prior to their arrival in the UK, as opposed to Group A, whose experience and native-speaker input in the early stages of learning was much less. This, however, appears to have implications in words that have no equivalent clusters in Greek, where Group A shows a slightly higher percentage of success compared to Group B, probably resulting from having received more accurate British English input. Speakers in both groups have begun learning English at approximately the same age, for a similar number of years, and with similar age of arrival and length of residence in the British-English-speaking community. Therefore, the difference in the degree of success between the two groups can be attributed to the differences in the quantity and quality of input received during the early stages of learning.

One implication of this study is that it proves the importance of the instruction of L2 pronunciation in early stages, especially when it takes place in a non-L2-speaking country. Greek-speaking schools could benefit from some routines used by English-speaking schools. For example, they could have at least some teachers of English that are native speakers of the language, or that have been assessed and found to have a near-native competence not only in knowledge of the grammatical structures, but in pronunciation accuracy as well. Importantly, this study was exploratory in nature, aiming to provide some preliminary observations on the English rhotic acquisition by native speakers of CyG in several contexts, and to offer possible lines per future investigations. Future research with specific focus on one or more contexts and with more tokens per context is required to provide a more complete picture of the production of the approximant by Greek or Greek-Cypriot learners of English.

Finally, the findings of this study should be interpreted with caution due to some methodological limitations. Firstly, language use patterns were obtained through self-reports, which is the most widely used but not necessarily reliable measure. In addition, the data were recorded in a controlled and not spontaneous environment, which makes speakers more conscious of their speech, resulting in better productions than might have been achieved in normal speech. However, structured elicitation was necessary in this study, to control for the contexts and vocalic or consonantal environments of the rhotic. Finally, the unequal number of speakers in each group, and of males and females within and across groups was not ideal, but unfortunately this was unavoidable due to the specific criteria needed to be fulfilled for participation, especially since students from Greek-speaking schools significantly outnumber students from English-speaking schools.

ACKNOWLEDGEMENTS

Firstly, I would like to thank the 28 Lancaster University students that provided the data in this study. I would also like to thank my MA supervisor Sam Kirkham for his valuable guidance and support, without which this study would not have been possible.

REFERENCES

Arvaniti, A. 1999a. Standard Modern Greek. *Journal of the International Phonetic Association* 29, 167–172.

Arvaniti, A. 1999b. Cypriot Greek. *Journal of the International Phonetic Association* 29, 173–178.

Arvaniti, A. 2001a. Cypriot Greek and the phonetics and phonology of geminates. *Modern Greek Dialects and Linguistics Theory* 1, 19–30.

Arvaniti, A. 2001b. Comparing the Phonetics of single and geminate consonants in Cypriot and Standard Greek. *4th International Conference on Greek Linguistics,* Thessaloniki, 37–44.

Arvaniti, A. 2007. Greek phonetics: The state of the art. *Journal of Greek Linguistics* 8, 97–208.

Arvaniti, A. and G. Tserdanelis. 2000. On the phonetics of geminates: Evidence from Cypriot Greek. *Proceedings of the 6th International Conference on Spoken Language Processing,* Beijing, China, 559–562.

Baltazani, M. 2005. Phonetic variability of the Greek rhotic sound. *Phonetics and Phonology in Iberia 5*.

Baltazani, M. 2009. Acoustic characterization of the Greek rhotic in clusters. *Proceedings of the 18th International Symposium of Theoretical and Applied Linguistics,* Thessaloniki, Greece, 87–95.

Baltazani, M. and K. Nicolaides. 2013. The many faces of /r/. In L. Speafico and A. Vietti (eds.) *Rhotics. New data and perspectives*. Bosen Bolsano: Bu Press, 125–144.

Baltazani, M. and K. Nicolaidis. 2012. Production of the Greek rhotic in initial and intervocalic position: An acoustic and electropalatographic study. *Selected papers of the 10th International Conference on Greek Linguistics*, 141–152.

Chan, A. Y. W. 2010. Advanced Cantonese ESL learners' production of English speech sounds: Problems and strategies. *System* 38, 316–328.

Cruttenden, A. and A. C. Gimson. 2014. *Gimson's pronunciation of English*. 8th edition. Abingdon: Routledge.

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 research*. Timonium, MD: York Press, 233–277.

Flege, J. E. 2003. Assessing constraints on second-language segmental production and perception. In A. Meyer and N. Schiller (eds.) *Phonetics and phonology in language comprehension and production, differences and similarities*. Berlin: Mouton de Gruyter, 319–355.

Flege, J. E. 2007. Language contact in bilingualism: Phonetic system interactions. In J. Cole and J. I. Hualde (eds.) *Laboratory Phonology* 9, Berlin: Mouton de Gruyter, 353–381.

Leather, J. and A. James. 1991. The acquisition of second language speech. *Studies in Second Language Acquisition* 13, 305–341.

Nicolaidis, K. and M. Baltazani, 2014. The Greek rhotic in /rC/ sequences: An acoustic and electropalatographic study. In N. Lavidas, T. Alexiou and A. M. Sougari (eds.) *Major trends in theoretical and applied linguistics*. Volume 1. De Gruyter, 157–176.

Nicolaidis, K. and M. Baltazani, 2015. An investigation of acoustic and articulatory variability during rhotic production in Greek. *Proceedings of the 11th International Conference on Greek Linguistics,* Rhodes, Greece.

Nicolaidis, K. and M. Baltazani. 2011. An electropalatographic and acoustic study of the Greek rhotic in /Cr/ clusters. *Proceedings of the 17th International Congress of Phonetic Sciences*, 1474–1478.

Olsen, M. K. 2012. The L2 acquisition of Spanish rhotics by L1 English speakers: The effect of L1 articulatory routines and phonetic context for allophonic variation. *Hispania* 95, 65–82.

Rose, M. 2010. Intervocalic tap and trill production in the acquisition of Spanish as a Second Language. *Studies in Hispanic and Lusophone Linguistics* 3, 379–420.

SUMMARY

L2 ACQUISITION AND PRODUCTION OF THE ENGLISH RHOTIC BY L1 GREEK-CYPRIOT SPEAKERS: THE EFFECT OF L1 ARTICULATORY ROUTINES AND PHONETIC CONTEXT

This study investigates the production of the English rhotic by Greek-Cypriot speakers, whose L1 typically involves a tap realisation. It also compares the productions of Greek-Cypriot learners that attended English-speaking schools with the productions of students that attended Greek-speaking schools during their secondary education. Participants were university students whose age of arrival in the UK was 17-21 years old, length of residence 1-4 years and age range 18-24 years. Six native speakers of English comprised the control group. Participants were recorded producing a Greek and an English wordlist with the rhotic in word-initial and intervocalic position, and in Cr and rC clusters. Manner of articulation, duration and formant frequencies were investigated. The results suggest that learners from English-speaking schools are more successful in the production of the English approximant than learners from Greek-speaking schools, although neither group reaches native-like values in all contexts in either duration or formant frequencies. Effects of the L1 phonetic system on L2 rhotic production are also found. This study provides insights on a subject that has received limited attention in the context of Cyprus, as well as a basis for future research that may lead to improvements in English language learning and teaching in Cyprus and other countries with similar phonetic inventories.

KEYWORDS: L2 rhotic production, SLM, phonetic category assimilation, phonetic category dissimilation, L1-L2 interaction, Greek-Cypriot learners of English.

APPENDIX

Table A1. Participant characteristics

Participant	Gender	YoS	Degree	Age	AOA	LOR	AOL
A1	Female	1	BSc Biomedical Sciences	19	18	1	9
A2	Male	4	MSc Data Science	23	20	4	8
A3	Male	4	MSc Biomedical Sciences	24	20	4	8
A4	Female	3	BSc Mathematics	21	18	3	9
A5	Female	3	BSc Mathematics	21	18	3	8
A6	Female	3	BSc Psychology	20	18	3	8
A7	Female	3	MSc Biological Sciences	21	18	3	9
A8	Female	3	BSc Biomedicine	20	18	3	7
A9	Female	3	BSc Economics	20	18	3	8
A10	Male	1	BEng Engineering	21	20	1	8
A11	Female	3	BSc Accounting and Finance	20	17	3	8
A12	Female	3	BSc Accounting and Finance	20	17	3	9
A13	Female	3	BSc Accounting and Finance	21	18	3	8
B1	Female	1	BSc Biomedical Sciences	18	17	1	7
B2	Female	1	BSc Biomedical Sciences	20	19	1	8
B3	Female	2	BSc Mathematics with Statistics	20	18	2	8
B4	Female	2	BSc Accounting and Management	20	19	2	9
B5	Female	2	BSc Mathematics	20	19	2	8
B6	Male	1	BSc Economics	21	21	1	10
B7	Male	1	BSc Accounting and Finance	20	19	1	10
B8	Female	3	BSc Accounting and Finance	21	18	3	8
B9	Female	3	Mathematics and Statistics	22	19	3	8
C1	Male	2	BSc Business Economics	19			
C2	Female	3	BA Ethics, Philosophy and Religion	21			
C3	Male	4	MA English Language and Literary Studies	22			
C4	Male	3	BA English Language	20			
C5	Male	2	BSc Physics	19			
C6	Male	2	BA English Literature and Philosophy	20			

Note: YoS: Year of Studies, AOA: Age of arrival in the UK, LOR: Length of residence in the UK, AOL: Age of learning of English as a foreign language in school.

ARTICLE INFO:
Original research article
Received: June 14, 2018
Revised: February 2, 2019
Accepted: February 13, 2019

UDC: 811.111'373.45:811.112.5
DOI: https://doi.org/10.18485/philologia.2018.16.16.4

THE ROLE OF ORTHOGRAPHY AND PHONEME INVENTORY IN DUTCH STUDENTS' SPEECH PERCEPTION IN THE EFL CLASSROOM

MARLISA HOMMEL[1]
University of Utrecht
The Netherlands

Cilj ovog rada je dvojak: da se odredi da li ortografija odmaže percepciji L2 kod učenika i da li, kao što PAM-L2 predviđa, fonemski inventar u L1 utiče na L2 govor holandskih srednjoškolaca. Najpre je sproveden pilot test kako bi se utvrdilo da li je teže opaziti 'nove' glasove, koji ne postoje u fonemskom inventaru holandskog, od 'istih' glasova. Nakon toga je sproveden još jedan test percepcije koji je ispitao da li ortografija otežava percepciju fonema ili grupa fonema koje su teške holandskim učenicima. U tom testu svi glasovi su se našli u frekventnim engleskim rečima i činili su izgovorene stimuluse, a pisane reči-mete su bile ili reči čiji pisani oblik odgovara auditorno percipiranim rečima, ili reči čiji se pisani oblik razlikuje od auditorno percipirane reči. Rezultati upućuju na to da ortografija otežava percepciju. Povrh toga, u skladu sa predviđanjima PAM-L2, 'nove' L2 glasove je teže percipirati od 'istih' glasova.

Ključne reči: perceptivno učenje, veliki fonemski inventar, holandski, engleski, ortografija.

1. INTRODUCTION

Learning a second or foreign language in school is often a laborious task. It comprises many aspects such as productive skills (speaking and writing), perceptive skills (listening and reading comprehension) and learning vocabulary and grammar rules. These language skills are related. Arguably the hardest to master are listening and speaking skills, because sounding native-like is strongly correlated with Age of Learning and these skills are online with little time for the learners to monitor

1 Contact information: m.p.hommel@uu.nl

their response (Flege *et al*. 1995). An adult language learner does not learn a foreign language in school by auditory input only, but also by orthographic input. This visual orthographic input can help or hinder the perception and production of L2 sounds. For example, Dutch does not have English /ð/ and /θ/ in its phoneme inventory, nor does it have the 'th' in its orthography and language learners will quickly realise that the English 'th' represents a new sound. In this instance, orthography may help the language learner. On the other hand, orthography can confuse the language learner. For example, the English written '-ough' is pronounced differently in 'tough' /tʌf/, 'though' /ðəʊ/ and 'through' /θruː/. Some languages have a high degree of grapheme-phoneme correspondence, making the orthography of these languages transparent. A study that looked at the role of orthography on the perceptual accuracy of speech sounds found that a language with transparent orthography has a facilitating effect on speech perception (Erdener and Burnham 2005). In the case of English /æ/-/e/[2] ('man'-'men'), Escudero *et al*. (2008) showed a facilitating effect of orthography of this English minimal pair for Dutch listeners. Many studies, however, found that orthography can hamper the "establishment of second language phonological categories" (Rafat 2011: iii, see also Bassetti *et al*. 2015 and Ziegler & Ferrand 1998). Silveira (2007), for instance, found a negative effect of English orthography on English pronunciation by ten Brazilian speakers. An opaque orthography or an "inconsistent spelling-to-sound mapping" in a language negatively influences word perception of that language (Ziegler & Ferrand 1998: 683). Aside from a few exceptions (such as /æ/-/e/), English has an opaque orthography which causes English children learning to read to lag behind in phoneme awareness compared to, for example, German or Dutch children as the latter languages have more transparent orthographies (Goswami *et al*. 2005). Orthography does not only influence phoneme awareness, but it is reciprocal: phoneme awareness also influences orthography (Patel *et al*. 2004). Because Dutch L1 speakers cannot readily apply their phoneme awareness to the opaque English orthography, it is expected that Dutch listeners are distracted by inconsistent sound-to-spelling mappings in English speech perception. This is the first question we will try to answer: does orthography play a negative role in L2 auditory speech perception?

Besides the role of orthography on L2 speech perception, we will look at the role of the L1 phoneme inventory on L2 speech perception. In order to do so, this paper evaluates predictions made by the Perceptual Assimilation Model (PAM). This speech perception model was proposed by Best (1995) and claims that naïve listeners perceive sound contrasts in different ways. The model has been adapted for adult L2 learners, PAM-L2 (Best & Tyler 2007). It postulates that the L1 filters how L2 sounds are perceived. When a non-native sound contrast is perceived as an existing sound contrast in the native phoneme inventory, native listeners discriminate this non-native contrast well (two-category assimilation). When a non-native sound contrast is perceived as belonging to only one sound in the native phoneme inventory, native listeners struggle to perceive this non-native contrast well. There are two ways that a non-native contrast is assimilated to a single phoneme: both sounds of a non-native sound contrast are perceived as good or poor exemplars of a single native phoneme (single-category

2 When referring to the DRESS vowel in this paper, RP symbol /e/ is used as opposed to GA /ɛ/ (Wells, 1982).

assimilation), or one non-native sound is considered a better exemplar of the native phoneme than the other (category goodness difference). In terms of perception, the two-category (TC) assimilation is perceived best by L2 listeners, followed by the category goodness (CG) and the single-category (SC) assimilation is most poorly perceived (TC > CG > SC) (Best 1995).

Applying PAM-L2 to this study, the model predicts that 'new' English sounds are poorer perceived by Dutch listeners than sounds that already exist in the Dutch phoneme inventory (Best & Tyler 2007). Two-category assimilation (for example L2 English /ɪ/-/i/ is perceived as L1 Dutch /ɪ/-/i/) is best perceived by L2 listeners, followed by one category assimilation (for example L2 English /æ/-/e/ are both perceived as L1 Dutch /ɛ/). In this paper, we do not distinguish between CG and SC and consider both categories as 'new' sounds and the TC as 'same' sounds because linguistic contexts in which phonemes are embedded can change vowel realisation to such extent that the listener may perceive it as either CG or SC (Strange *et al.* 2001). The second question we will try to answer is: do Dutch students perceive English 'same' sounds better than 'new' sounds?

2. PILOT PERCEPTION TASK TO EVALUATE PAM-L2

In a previous study (Hommel 2017), an auditory speech perception pilot was conducted to see which English phonemes are difficult for Dutch students to perceive. Out of an identification task with 109 different phonemes and phoneme clusters, thirty-eight speech sounds that were poorly perceived by Dutch listeners were identified (see Table 1 below). A new test was created based on the results of the pilot, which consisted solely of these 38 constituents. This test involved not only auditory words but also written words (whereas the pilot only involved written letters representing sounds). Students would *hear* a word, e.g. 'p**u**sh' and saw four options *written* on a screen: 'c**oo**k', 'c**o**pe', 'c**o**w' and 'c**ou**p'. Students had to choose what sound they had heard in the auditory target by clicking on the word with underlining in the same position as the corresponding target sound, which could be in any position (onset, nucleus or coda). The correct answer in this example was 'cook' because both 'push' and 'cook' have the /ʊ/-sound in the underlined position. Had the response options been heard instead of written, it would have asked too much of the auditory working memory of the student. In addition, the test was deemed too easy if the auditory word would be the same as the written target answer (e.g. if 'p**u**sh' was heard and 'p**u**sh' would be the written answer).

Onset	Nucleus	Coda
/j/	/ɒ/	/θ/
/v/	/ʊ/	/g/
/z/	/əʊ/	/t/
/d/	/ʌ/	/dʒ/
/l/	/ɪ/	/z/
/g/	/e/	/s/
/f/	/ɔ/	/f/
/dʒ/	/i/	/tʃ/
/s/	/eɪ/	/k/
/ð/	/u/	/v/
/r/	/æ/	/d/
/θ/		/ʃ/
/b/		/ʒ/
		/ð/

Table 1. Difficult English sounds by phonological position

We looked at students' average score per speech sound in this task to compare 'new' with 'old' sounds. However, it is debatable which English sound is considered 'new' for Dutch listeners. PAM uses the International Phonetic Alphabet (IPA) as a basis for determining a 'new' or 'same' sound, which can be unreliable (Rochet 1995). Therefore, in addition to using IPA to decide whether an English sound is 'new' or not for Dutch native listeners, we also consulted English and Dutch vowel and consonant charts as phoneme realisations can differ despite sharing the same IPA symbol (see e.g. Levy & Law 2010). These charts represent the acoustic realisations of phonemes. Vowel charts show that Dutch /ɛ/ differs from English /e/ (Gussenhoven 1999: 76 for the Dutch chart and Roach 2004: 242 for the BrE chart). They also show that Dutch /ɔ/ has a very different realisation than the English /ɔ:/ which is why this nucleus is also considered 'new' to Dutch L1 speakers. Concerning consonants, /ʒ/ is an uncommon sound in Dutch (e.g. in the French loan word *baggage* [baˈχaːʒə] 'luggage'). However, /ʒ/ in coda position is rare in English too and occurs, similar to Dutch, in French loan words (e.g. beige [beɪʒ]). When comparing the Dutch and English IPA charts and adding English /ɔ:/, eleven of the 29 constituents have no direct Dutch counterpart, namely consonant (clusters) /tʃ/, /dʒ/, /θ/ and /ð/ and nuclei /ɒ/, /ʊ/, /əʊ/, /ʌ/, /æ/, /e/ and /ɔ:/ (Mees & Collins 2003). Linguistic contexts in which phonemes are embedded can change vowel realisation to such extent that the listener will perceive it as either CG or SC (Strange *et al.* 2001).

It should be noted that "neither very abstract phonological descriptions of phoneme inventories nor acoustic comparisons of specific realizations of phoneme categories will be adequate in predicting cross-language perceptual similarities"; whether an L2 sound is perceived as 'same' or 'new' changes depending on the context that the sound is in (ibid: 1703). The pilot contained no orthography, making it easier to determine whether a sound is perceived as 'same' or 'new'.

3. METHOD

For evaluating PAM-L2, students' perception scores on the pilot perception task were used. Out of the 38 difficult sounds (see Table 1), nine sounds occur in both onset and coda position, leaving 29 different constituents. In order to establish whether Dutch listeners perceive 'new' sounds (/tʃ/, /dʒ/, /θ/, /ð/ and nuclei /ɒ/, /ʊ/, /əʊ/, /ʌ/, /æ/, /e/ and /ɔ:/) more poorly than 'same' sounds (/l/, /f/, /v/, /s/, /z/, /ʃ/, /ʒ/, /r/, /b/, /j/, /t/, /d/, /k/, /g/, and nuclei /ɪ/, /i:/, /eɪ/ and /u:/), perception scores of 'new' sounds in the pilot were compared to 'same' sounds using an independent t-test. Students' perception scores from the pilot were used (N = 50, M = 19 years, SD = 2.1 years), see Appendix A for mean scores per constituent.

For evaluating whether English orthography plays a negative role in L2 auditory speech perception, another speech perception task was conducted. We call this the main test as the number of participants was much higher (N = 125) than the pilot described in section two (N = 50). Students' average perception scores on sounds with a difficult orthography were compared to test scores on sounds with an easy orthography. Words were considered to have a difficult or opaque orthography where the orthography of the spoken word differed from that of the written word being the correct response. For example, target nucleus /ʊ/ was seen as difficult because the /ʊ/ in 'p**u**sh' differed in orthography from the correct answer 'c**oo**k' (see Table 2). Next, words were considered to have a difficult or opaque orthography where the target answer was similar to a distractor. For example, target nucleus /ʌ/ was seen as difficult because the /ʌ/ in correct 'b**u**t' is similar in orthography to the incorrect foil 'p**u**ll' (see table 3). Independent t-tests were performed to see whether orthography influences listeners' speech perception. Students' perception scores on the main test were used (N = 125, M = 19 years, SD = 1.3 year).

The main test was not used for evaluating PAM-L2 as many of the new sounds also contained target sounds with different orthography (namely /dʒ/, /ʊ/ and /ɔ:/, see table 2) or good distractors with similar orthography (namely /θ/, /əʊ/, /ʌ/ and /æ/, see table 3), which is a confounding factor. If new sounds were to score significantly lower than familiar/same sounds, then this could be due to orthography rather than new vs. same sounds. The pilot contains little to no orthography, making it more suitable to evaluate PAM-L2. The 'new' sound coda /ð/ did not occur in the pilot.

4. RESULTS

4.1 ORTHOGRAPHY

In the main test, six out of thirty-eight target words contained a different orthography to the spoken word (see Table 2). These six words (M = 45%, SD = 28.4%) scored significantly lower compared to the target sounds with similar orthography (N = 32, M = 71%, SD = 24.6%), $t(36) = -2.31$, $p = .03$. The effect size ($d = 0.97$) signifies a large difference (Becker 2000). These six target sounds that contained a different orthography were excluded. Next, we looked at foils that contained similar orthography to the spoken words (i.e., good distractors) and they were compared to average perception scores on the remaining sounds (see table 3). Results show that foils with similar orthography confused the listeners more (N = 11, M = 56%, SD = 28.6%) than others (N = 21, M = 78%, SD = 18.8%), $t(30) = -2.62$, $p = .01$. The effect size ($d = .91$) signifies a large difference (Becker 2000).

Because the result suggests that orthography hampers perception and there is little to no orthography in the pilot, one would expect the scores in the pilot to be higher than the scores in the main test. However, this is not the case. Average scores in perception seem very comparable (71% in the pilot vs. 74% in the main test in onsets, 62% vs. 57% in nuclei and 68% vs. 67% in codas respectively).

Target sound	Auditory word	Written target
/ʊ/	Push	Cook
/ɔ:/	Thought	Paw
/i:/	Deal	Keep
/-dʒ/	Age	Ju**dge**
/dʒ-/	Gym	Duke
/-s/	Face	Mess

Table 2. English auditory words with target answers with a different orthography

Target sound	Written target	Foil
/-θ/	Oath	Clothe
/əʊ/	Poke	Pot
/ʌ/	But	Pull
/ɪ/	Bit	Bite
/-z/	His	Kiss
/eɪ/	Cape	Scalp
/u:/	Boot	Book
/æ/	Cap	Cape
/ð-/	Though	Thief
/θ-/	Theme	These
/-ð/	Sunbathe	Beneath

Table 3. English written target answers with good distractors

4.2 NEW VS. SAME SOUNDS

In order to establish whether Dutch listeners perceive 'new' sounds more poorly than 'same' sounds, perception scores on the pilot were compared. Results show that 'new' sounds (N = 12, M = 56%, SD = 22.6%) were harder to perceive than 'same' sounds (N = 22, M = 74%, SD = 18.7%), $t(32) = -2.54$, $p = .02$. The effect size ($d = .87$) signifies a large difference (Becker 2000).

5. DISCUSSION AND CONCLUSION

In this paper we looked at two hypotheses: (1) that opaque orthography impedes perception and (2) that, in accordance with many other studies (e.g. Kartushina *et al*. 2015; Levy & Law 2010 and Tyler *et al*, 2014), 'new' sounds are more difficult to perceive than 'same' sounds, and that this still holds for a classroom setting. Both hypotheses were confirmed. Results suggest that orthography impedes perception but that linguistic context (i.e. using stimuli in words instead of isolated stimuli) may facilitate perception. Listening to L2 phonemes embedded in a linguistic context (in the main test this context was a high frequency word) helps the language learner in forming new phoneme categories (Strange *et al*. 2001). Phonemes are pronounced differently depending on the context due to co-articulation, and the context is one of the factors

that determine how difficult the sound is to perceive for a language learner (ibid). Therefore, the ecological validity of the task (i.e. using high-frequency words) seems to justify the use of phonemes in a linguistic context as opposed to a task where these phonemes only occur in isolation.

When a sound is absent in the Dutch phoneme inventory (a 'new' sound), it is difficult for Dutch listeners to perceive this sound well. When looking at the perception score of each individual phoneme, however, not all 'new' sounds were hard to perceive. For example, the 'new' sounds /æ/ and /e/ were not poorly perceived and the 'new' sound /ʊ/ was not perceived worse than 'same' (but still differing acoustically) sound /u:/. It is possible that speech perception is influenced by statistical learning, where new L2 sounds which are more frequent are learned earlier and/or better as they are more salient in the L2 (Mines *et al.* 1978). Because certain contrasts are more frequent than others, it is especially important for language learners to correctly discern the most common contrasts. For example, English /æ/-/e/ is more frequent than the minimal pair /ʊ/-/u:/, making the former contrast more salient. Not surprisingly, Dutch students correctly identified both /æ/ and /e/ better than /ʊ/ and /u:/. Similarly, Dutch students perceived both /ʊ/ and /u:/ better than less common /θ/ and /ð/, perhaps due to salience or statistical learning. Phoneme /θ/ is more frequent and less marked than voiced /ð/ which could explain why /θ/ was perceived better in both onset (/θ/ 62% vs. /ð/ 27%) and coda position (/θ/ 65% vs. /ð/ 6%).

Results found here are tentative. To have more robust evidence that orthography influences listeners' speech perception, the same stimuli should be tested with auditory stimuli that both *differ in* and *share the same* orthography with the written target answer, not with either or. In addition, another confounding factor is that 'new' sounds are perceptually more difficult than 'same' sounds. Future research looking at the role of orthography in L2 speech perception should take into account the factor of 'new' and 'same' sounds.

ACKNOWLEDGMENTS

This work is part of the Doctoral Grant for Teachers with project number 023.007.008, which is financed by the Netherlands Organisation for Scientific Research (NWO). I would like to thank my supervisors René Kager and Willemijn Heeren for their time and their valuable suggestions.

REFERENCES

Bassetti, B., P. Escudero & R. Hayes-Harb. 2015. Second Language Phonology at the Interface between Acoustic and Orthographic Input. *Applied psycholinguistics* 36 (1), 1–6.

Becker, L. 2000. Effect Size Calculators. [Internet]. Available at: http://www.uccs.edu/lbecker/index.html [30.01.2018].

Best, C. T. 1995. A Direct Realist View of Cross-Language Speech Perception. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Research*. York: Timonium, 171–204.

Best, C. T. & M. D. Tyler. 2007. Nonnative and Second-language Speech Perception: Commonalities and Complementarities. In O. S. Bohn & M. J. Munro (eds.) *Language Experience in Second Language Speech learning: In honor of James Emil Flege*. Amsterdam: John Benjamins Publishing, 13–34.

Cutler, A. *et al*. 2004. Patterns of English Phoneme Confusions by Native and Non-native Listeners. *JASA* 116 (6), 3668–3678.

Erdener, V. D. & D. K. Burnham. 2005. The Role of Audiovisual Speech and Orthographic Information in Nonnative Speech Production. *Language Learning* 55, 191–228.

Escudero, P., R. Hayes-Harb & H. Mitterer. 2008. Novel L2 Words and Asymmetric Lexical Access. *Journal of Phonetics* 36, 345–360.

Escudero, P., E. Simon & H. Mitterer. 2012. The Perception of English Front Vowels by North Holland and Flemish Listeners: Acoustic Similarity Predicts and Explains Cross-linguistic and L2 Perception. *Journal of Phonetics* 40 (2), 280–288.

Flege, J. E. 1992. The Intelligibility of English Vowels Spoken by British and Dutch Talkers. *Intelligibility in speech disorders: Theory, measurement, and management* 1, 157–232.

Flege, J. E. & W. Eefting. 1987. Cross-language Switching in Stop Consonant Perception and Production by Dutch Speakers of English. *Speech Communication* 6 (3), 185–202.

Flege, J. E., M. J. Munro & I. R. MacKay. 1995. Factors Affecting Strength of Perceived Foreign Accent in a Second Language. *JASA* 97 (5), 3125–3134.

Goswami, U., J. C. Ziegler & U. Richardson. 2005. The Effects of Spelling Consistency on Phonological Awareness: A Comparison of English and German. *Journal of experimental child psychology* 92 (4), 345–365.

Gussenhoven, C. 1999. Dutch. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Hommel, M. 2017. Speech Perception Training as a Serious Game in the EFL Classroom. *Proceedings of the Seventh ISCA workshop on Speech and Language Technology in Education Stockholm*, 177–180.

Johnson, K. & M. Babel. 2010. On the Perceptual Basis of Distinctive Features: Evidence from the Perception of Fricatives by Dutch and English Speakers. *Journal of phonetics* 38 (1), 127–136.

Kartushina, N. *et al*. 2015. The Effect of Phonetic Production Training with Visual Feedback on the Perception and Production of Foreign Speech Sounds. *JASA* 138 (2), 817–832.

Levy, E. S. & F. F. Law. 2010. Production of French Vowels by American-English Learners of French: Language Experience, Consonantal Context, and the Perception-production Relationship. *JASA* 128 (3), 1290–1305.

Major, R. C. 2001. *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. New York: Lawrence Erlbaum.

Mees, I. & B. Collins 2003. *The Phonetics of English and Dutch*. Leiden: Brill.

Mines, M. A., B. F. Hanson & J. E. Shoup 1978. Frequency of Occurrence of Phonemes in Conversational English. *Language and speech* 21 (3), 221–241.

Patel, T. K., M. J. Snowling & P. F. de Jong. 2004. A Cross-linguistic Comparison of Children Learning to Read in English and Dutch. *Journal of Educational Psychology* 96 (4), 785.

Rafat, Y. 2011. *Orthography-induced Transfer in the Production of Novice Adult English-speaking Learners of Spanish*. Dissertation. University of Toronto. [Internet]. Available at: https://tspace-library-utoronto-ca.proxy.library.uu.nl/handle/1807/31909 [12.07.2018].

Roach, P. 2004. British English: Received Pronunciation. *Journal of the International Phonetic Association* 34 (2), 239–245.

Rochet, B. L. 1995. Perception and Production of Second-Language Speech Sounds by Adults. In W. Strange (ed.) *Speech Perception and Linguistic Experience.* Timonium, MD: York Press, 379–410.

Silveira, R. 2007. The Role of Task-type and Orthography in the Production of Word-final Consonants. *Revista de Estudos da Linguagem* 15 (1), 143–176.

Simon, E. 2010. Phonological Transfer of Voicing and Devoicing Rules: Evidence from L1 Dutch and L2 English Conversational Speech. *Language Sciences* 32 (1), 63–86.

Strange, W. *et al*. 2001. Effects of Consonantal Context on Perceptual Assimilation of American English Vowels by Japanese Listeners. *JASA* 109 (4), 1691–1704.

Tyler, M. D. *et al*. 2014. Perceptual Assimilation and Discrimination of Non-native Vowel Contrasts. *Phonetica* 71 (1), 4–21.

Wells, J. C. 1982. *Accents of English*, *The British Isles*, vol. 2. Cambridge: Cambridge University Press.

Wester, F., D. Gilbers & W. Lowie. 2007. Substitution of Dental Fricatives in English by Dutch L2 Speakers. *Language Sciences* 29, 477–491.

Ziegler, J. C. & L. Ferrand. 1998. Orthography Shapes the Perception of Speech: The Consistency Effect in Auditory Word Recognition. *Psychonomic Bulletin & Review* 5 (4), 683–689.

SUMMARY

THE ROLE OF ORTHOGRAPHY AND PHONEME INVENTORY IN DUTCH STUDENTS' SPEECH PERCEPTION IN THE EFL CLASSROOM

The goal of this paper is twofold: to determine whether orthography hampers students' L2 speech perception and whether, as PAM-L2 predicts, the L1 phoneme inventory influences L2 speech perception of Dutch secondary school students. First, a pilot perception test was administered to see if 'new' sounds that do not exist in the Dutch phoneme inventory are harder to perceive than 'same' sounds. Next, another perception test was created to look at whether orthography hampers Dutch students' perception of -for Dutch listeners- difficult English phonemes or phoneme clusters. In that perception test, all phoneme (cluster)s were embedded in common English words as spoken stimuli, and the written target words were sometimes words with an orthography that corresponded with the auditorily perceived word and sometimes with an orthography that differed from the auditorily perceived word. Results indicate that orthography impedes perception. Furthermore, in line with the predictions of PAM-L2, 'new' L2 sounds are harder to perceive than 'same' sounds.

KEYWORDS: perceptual learning, large phoneme inventory, Dutch, English, orthography.

APPENDIX

Appendix A: English constituents correctly perceived on the pilot P and main test T1 respectively

Onset P T1	Nucleus P T1	Coda P T1
/j/ 56% 96%	/ɒ/ 34% 39%	/θ/ 76% 65%
/v/ 66% 67%	/ʊ/ 90% 36%	/g/ 72% 84%
/z/ 90% 93%	/əʊ/ 44% 81%	/t/ 70% 86%
/d/ 82% 81%	/ʌ/ 36% 24%	/dʒ/ 44% 75%
/l/ 90% 80%	/ɪ/ 82% 98%	/z/ 46% 65%
/g/ 78% 99%	/e/ 78% 59%	/s/ 78% 35%
/f/ 56% 50%	/ɔ:/ 50% 13%	/f/ 46% 54%
/dʒ/ 90% 25%	/i:/ 88% 84%	/tʃ/ 64% 76%
/s/ 82% 93%	/eɪ/ 86% 64%	/k/ 88% 94%
/ð/ 24% 27%	/u:/ 24% 39%	/v/ 52% 86%
/r/ 74% 95%	/æ/ 72% 86%	/d/ 86% 80%
/θ/ 36% 62%	**62% 57%**	/ʃ/ 94% 93%
/b/ 98% 94%		/ʒ/ - 42%
71% 74%		/ð/ - 6%
		68% 67%

ARTICLE INFO:
Original research article
Received: September 14, 2018
Revised: February 8, 2019
Accepted: February 12, 2019

UDC: 811.111'342
811.111'344
DOI: https://doi.org/10.18485/philologia.2018.16.16.5

CONTEMPORARY PHONOLOGICAL THEORY AND COMMON PRONUNCIATION ISSUES

VIŠNJA JOSIPOVIĆ SMOJVER[1]
Faculty of Humanities and Social Sciences
University of Zagreb
Zagreb, Croatia

U ovom radu razmatra se uloga suvremene fonološke teorije u poučavanju engleske fonetike na sveučilišnoj razini. Pobliže, pokazuje se kako, nasuprot uvriježenom popularnom vjerovanju, nastava engleske fonetike za studente engleskoga jezika nema za cilj da se studente uči da engleski izgovaraju „kultivirano", što god to u današnje doba podrazumijevalo. Budući da sveučilišna izobrazba budućih stručnjaka za engleski jezik zahtijeva stjecanje uvida u prirodu izgovornih pojava, u ovom se radu pokazuje kako fonološka teorija može rasvijetliti neke praktične izgovorne probleme. To se čini pozivanjem na teorijske okvire poput grkljanske fonologije, teorije otvora, artikulacijske teorije te optimalnosne teorije, da bi se pokazalo kako podizanje razine fonološke osviješćenosti o izgovornim problemima može razviti senzibilitet za važna izgovorna pitanja.

Ključne riječi: fonološka teorija, izgovorne pojave, poučavanje engleske fonetike, fonološka osviješćenost.

1. INTRODUCTION

The present paper addresses two basic kinds of pronunciation issues that contemporary phonological theory can throw light on. One of them concerns a rather practical issue in mastering English pronunciation, namely, the apparent intelligibility of the English lyrics of popular songs and the way in which some theoretical phonological frameworks can account for this. The other one includes theoretical phonological issues which are generally useful in developing sensitivity to important English pronunciation features. The former is going to be analysed in connection with three popular songs

1 Contact information: visnja.josipovic@zg.t-com.hr

(*The Road to Mandalay*, *Budapest*, and *Sign of the Times*) and with reference to the broad framework of Feature Geometry, in particular, the dimensional theory of laryngeal features, as in Iverson and Ahn (2007); the account of lenition processes as given by Gurevich (2011) or Ernestus (2011); Aperture Theory (Grijzenhout 1996, 2011), and Optimality Theory (McCarthy 2007, 2008).

As for the general theoretical issues that are of some interest in teaching English phonetics at university level, they are twofold. They include pronunciation features related to the typological prosodic differences between English and the students' first language, as well as some amusing, but nonetheless useful, theoretical phonological insights contributing to raising phonological awarenes, such as ludlings and expletive infixation.

2. PRACTICAL ISSUES

Concerning the practical issues related to English pronunciation, reference is going to be made to three popular songs, all of them number-one hits from the world's top charts. The first is the Robbie Williams 2001 hit, *The Road to Mandalay*. What initially drew the attention of the present paper's author to its lyrics was an informal survey carried out by Mateusz-Milan Stanojević (personal communication) among a group of Croats, including students and teachers from the Department of English, University of Zagreb. When asked to write down the refrain of the song, each of them responded by '***P****a ram* ***p****am* ***p****am* ***p****am* ***p****am* ***p****am* ***p****am*', as opposed to the original refrain text of the lyrics offered on YouTube (https://genius.com), '***B****om* ***b****om* ***b****om* ***b****a da dup* ***b****om* ***b****om*'.[2] When the author of the present paper, who without hesitation would have responded in exactly the same way, checked the original lyrics in disbelief, there followed another shock: after '*Save me from drowning in the sea*' what followed was ***'beat me up*** *on the beach*', rather than ***'pick me up*** *on the beach*', which any speaker of Croatian would probably have heard!

To account for this kind of misperception, contemporary phonological theory offers some interesting insights into the difference between aspiration languages and voicing languages within the broader framework of laryngeal phonology, which ultimately can be placed into a yet wider framework of feature geometry. In particular, reference should be made to the so-called dimensional theory of laryngeal features, as proposed by Iverson and Ahn (2007). These two authors account nicely for the true nature of the difference between aspiration languages and voicing languages.

The gist of their theory is that in aspiration languages, such as English and the majority of Germanic languages, voicing is not present in the phonological representation. That is, in these languages segments are unspecified for voicing. Rather, they are specified either for the feature 'spread glottis' [sg] or 'constricted glottis' [cg], which are seen as privative, unary, i.e., non-binary features. Under this approach voicing is seen as redundant in aspiration languages: in sonorants it is spontaneous, whereas in voiced obstruents, which are not specified as [sg], it is passive.

On the other hand, the category of voicing languages, for which Iverson and Ahn (2007) take French as an outstanding example, obviously includes languages like

2 https://genius.com/Robbie-williams-the-road-to-mandalay-lyrics [June 12, 2018]

Croatian and Serbian as well. In these languages segments are specified for voicing, but not for the other two laryngeal features, [sg] and [cg].

Although it is indisputable that the contrast between voiced and voiceless consonants does exist in English, in the dimensional theory this is interpreted as a surface phonetic contrast on the post-lexical level, functioning to enhance the truly phonologically relevant contrast between the gestures of opening and closing the glottis. Phonetically, the feature of spread glottis is realised as audible breath, i.e., aspiration in the right context, where the contrast needs to be enhanced, which is the foot-initial position in English (e.g. *'a tissue'*, as opposed to *'at issue'*, to quote the famous phonology-textbook example).

Bearing all this in mind, let us return to our initial issue of the misperception of voicing in English by speakers of Croatian and explain why it happens. In short, the two languages are typologically different with respect to the status of voicing. In English, as opposed to Croatian, voicing is a redundant feature, which is only passive in voiced obstruents, such as the /b/ in the examples from the Robbie Williams song referred to above. The phonetic manifestation of this difference between the two languages under consideration is the misleading difference in the voice onset time.

Taking all this into consideration, it is no surprise then that Jenkins (2000) includes aspiration among features of English pronunciation that belong to the so-called *Lingua Franca Core*, i.e., those that are likely to be associated with the international unintelligibility of English. This also explains the numerous jokes and anecdotes in Croatian based on situations of misunderstandings between Germans and Croats, like the one involving the German word *'danke'* ('thank you') misperceived as *'tanke'* (meaning 'thin' in Croatian).

Perhaps another point to note in connection with the refrain of *The Road to Mandalay* concerns the flap, indicatively spelled as ***d*** in the English version (*'...ba **d**a **d**up...')* and as **r** in its Croatian counterpart. It is the underlying [cg] coronal stop, which is phonetically not only passively voiced, but also lenited (weakened) in the lenis intervocalic surroundings. However, this issue of lenition brings us to the next example of song lyrics misperception discussed in the present context.

In George Ezra's 2014 hit *Budapest* (https://genius.com)[3] there is another example of the same kind of lenition when the /t/ of 'it' in *'I'd leave **it** all'* and *'I'd lose **it** all'* is typically perceived by Croatian speakers as /d/, within the wider context of the stretches misperceived as [aɪ ˌdɪvɪ ˈdɔ:] and [aɪ ˌdu:zɪ ˈdɔ:]. Here one can observe extreme cases of lenition affecting the coronal lateral /l/, a type of segment which is really special in phonological theory in that there are apparently more controversial and unresolved issues about it than for any other segment (cf. Yip 2011).

All the major theoretical frameworks which have provided some interesting insights into the nature of this kind of lenition in one way or another treat sonority as an inherently gradual feature, as opposed to the majority of other segmental features, which are treated as non-gradual, i.e., either binary or monovalent. One of them is Aperture Theory, as set out by Grijzenhout (1996). The way in which the cases of lenition under consideration are analysed in this theory is to operate with nodes of aperture

3 https://genius.com/George-ezra-budapest-lyrics [June 12, 2018].

(ranging from A_0 to A_{max}) within the feature geometry tree to account for potentially gradual features, such as sonority or nasality. Lenition processes like those affecting the English lateral /l/ in the example at hand, just like the opposite fortition processes are then interpreted as delinking nodes representing particular degrees of opening, or, respectively, segments docking onto nodes representing certain aperture degrees.

Likewise, authors like Gurevich (2011), for example, or Ernestus (2011), explicitly operate with a hierarchy of lenition expressed on a scale ranging from geminates at the one end to zero, corresponding to the loss of a segment. This latter extreme would correspond to the complete loss of the /l/ in the examples at hand, [aɪ ˌdɪvɪ ˈdɔ:] and [aɪ ˌdu:zɪ ˈdɔ:], where the /l/ of *'leave'* and *'lose'* are completely lost, whereas the degree of the lenition of the final /l/ of 'all' is disputable and variable, depending on the listener: it can be anything ranging from a 'dark', velarised [ɬ], through the fully vocalised one, [ɯ], down to the complete loss, i.e., zero, as in *'leave'* and *'lose'*.

Finally, Optimality Theory with Candidate Chains, OT-CC, as in McCarthy (2007) has its own way of dealing with what is perceived as lenition processes – postulating a level of candidate chains in which each candidate is exactly by one grade more harmonic than the previous one.

The last song which will be referred to here is Harry Styles' 2017 number one in the UK, *Sign of the Times* (https://genius.com)[4]. The focus will be on the reason that someone, especially a speaker of a pitch-accent language, is likely to interpret *'the bullets'* as sung in the example at hand as a separate word, something like *'topolits'*.

What elegantly explains this are two OT constraints, working to the same effect here. They are most blatantly violated here by the unfortunate match of the melody and the text, when the melody goes sky-high on the definite article. One of the constraints under consideration is known as ROOTING (cf. McCarthy 2008). It generally requires the prosodic prominence of lexical morphemes, i.e. roots. In the analysis of tone languages there is a similar constraint, in fact somewhat broader in scope, requiring the linking of H tones to prosodically or morphologically prominent positions. Needless to say, even though English is not a tone language, the extremely high pitch on the definite article for a speaker of a pitch-accent language definitely equals H tone, which in its turn equals accent, and a sequence perceived like /'topolits/ equals a prosodic and morphological word!

3. GENERAL PRONUNCIATION ISSUES

In terms of other phonological insights useful for teaching English phonetics to speakers of pitch-accent languages, a special place belongs to the OT account of *Neoštokavian* pitch-accent prosody as given by Zec (1999). Although as a speaker of Serbian the author naturally refers to Serbian, it should be noted that this kind of theoretical account is applicable to Croatian as well, as these two cognate languages belong to the same category in terms of prosodic typology.

In this context we may return to the point made in connection with the unintelligibility of *'the bullets'* in *Signs of the Time*. What throws additional light on the

4 https://genius.com/Harry-styles-sign-of-the-times-lyrics [June 12, 2018]

issue is one of the crucial ideas of the account of *Neoštokavian* provided by Zec (1999), namely, that of tonal feet. When it receives a high tone, a syllable acquires the status of a tonal foot (as opposed to a metrical foot), and within the framework of this very insightful account, *Neoštokavian* prosody is explained in terms of a complex interaction of OT constraints characterising trochaic rhythm and pitch-accent prosody.

This kind of approach again clears up a number of practical pronunciation issues in the languages concerned, such as the misperception of song lyrics in L_1, analogous to those discussed above in connection with teaching English phonetics. It also helps to understand the awkward and unnaturally-sounding accentual patterns in L_1 which are commonly heard in the media, as well as patterns behind L_1 prosody acquisition, as discussed in Josipović Smojver (2003; 2017), to name a few. Needless to say, specialists in a foreign language are expected to have some basic insights about their own language, as well as to be able to describe their native language in the first place.

Finally, mention should be made at this point of some amusing theoretical insights provided by contemporary phonological theory, which can teach us about various English pronunciation phenomena. Among them are the so-called ludlings (also known as language games, secret languages). There are lots of various ludlings, based on different languages, including English (e.g. *Pig Latin*, *Ubbi-Dubbi*, *Schmanguage*, *Shizzolation*, *Spaka, Homeric*) and Croatian (e.g. *Jepezipik,* the Croatian version of the internationally known *Parrot Language*), as well as those based on other, better or lesser known world languages.[5] The morphoprosodic templates used in the word formations of these secret languages can teach a student of English phonetics a lot about syllable structure – both universal and language-specific aspects of it. Ludlings are also instructive in connection with the role of prosodic domains for expressing pronunciation rules and regularities. A case in point would be hypocoristic formation, which generally follows morphoprosodic templates comparable to those of ludlings in both English and Croatian.

The prosodic domains commonly used in ludling word formation include not only the syllable with its constituents,which are important, for example, in *Pig Latin* and *Jepezipik,* but also the foot, mentioned above in connection with the English aspiration rule. It is precisely the foot that plays the crucial role in *Homeric,* described by Vaux (2011: 741). This brings us to the other example of how contemporary phonological theory can throw light on pronunciation phenomena in an amusing way. Just like the language game of *Homeric*, the English phenomenon of expletive infixation relies on the foot as the relevant analytical entity in phonology, following the principle of inserting the 'dirty-word' infix consisting of a foot, i.e. a sequence of a stressed and an unstressed syllable. Typical expletive infixes following this template would be *bloody*, or *fucking*, for example, which are inserted between the feet of the original, starting-point words of 'regular' English, resulting in expletive forms like *Ala-bloody-bama* for *Alabama*. (cf. McCarthy 1982). An analogous Croatian procedure is discussed in Josipović Smojver (2017) with reference to the example *Erasmus - vražji - ugovor*.

5 Elaboration of the topic, as well as detailed references are given in Josipović Smojver (2017).

4. CONCLUSION

By way of conclusion, it turns out that constant references to contemporary phonological issues are necessary in teaching English at university level. Nowadays there are influential linguistic theories and phonological models that no longer even see phonology and phonetics as separate components, cases in point being Articulatory Theory, as in Browman and Goldstein (1992) or Zsiga (2018), as well as feature theories focusing on the gradient nature of certain phonological features, such as Ernestus (2011) or Gurevich (2011).

It seems indisputable that phonological insights into the nature of pronunciation phenomena develop sensitivity to pronunciation issues. Provided that such an attitude is encouraged in phonetics teaching at university level, the ultimate outcome is going to be not only a more cultivated English pronunciation among future specialists in the English language, but also, notably, a greater degree of effectiveness in cultivating others' English pronunciation in the broadest sense.

REFERENCES

Browman, C. and L. Goldstein. 1992. Articulatory Theory: An overview. *Phonetica* 49, 155–180.

Ernestus, M. 2011. Gradience and Categoricality. In M. van Oostendorp, C. Ewen, E. Hume and K. Rice (eds.) *The Blackwell Companion to Phonology. Vol. IV, Phonological Interfaces*. Chichester, UK: Wiley-Blackwell, 2115–2136.

Grijzenhout, J. 1996. Consonant weakening processes and Aperture Theory. *Toronto Working Papers in Linguistics* 15, 105–124.

Grijzenhout, J. 2011. Consonant Mutation. In M. van Oostendorp, C. Ewen, E. Hume and K. Rice (eds.) *The Blackwell Companion to Phonology, Vol. III*, *Phonological Processes*. Chichester, UK: Wiley-Blackwell, 1537–1558.

Gurevich, N. 2011. Lenition. In M. van Oostendorp, C. Ewen, E. Hume and K. Rice (eds.) *The Blackwell Companion to Phonology, Vol. III*, *Phonological Processes*. Chichester, UK: Wiley-Blackwell, 1559–1575.

Iverson, G. K. and S. C. Ahn. 2007. English Voicing in Dimensional Theory. *Language Sciences*, May, 29 (2.3), 247–269. doi: 10.1016/j.langsci.2006.12.012.

Jenkins, J. 2000. *The Phonology of English as an International Language: New Models, New Goals*. Oxford: Oxford University Press.

Josipović Smojver, V. 2003. Ema, Nina i Emanina: analiza slučaja blizanačkog govora. *Govor* XX, 1–2, 169–179.

Josipović Smojver, V. 2017. *Suvremene fonološke teorije*. Zagreb: Ibis grafika.

McCarthy, J. 1982. Prosodic Structure and Expletive Infixation. *Language* 58 (3), 574–590.

McCarthy, J. 2007. *Hidden Generalizations: Phonological opacity in Optimality Theory*. London: Equinox.

McCarthy, J. 2008. *Doing Optimality Theory: Applying Theory to Data*. Maiden, MA: Blackwell.

Vaux, B. 2011. Language Games. In J. Goldsmith, J. Riggle i A. C. L. Yu (eds.) *The Handbook of Phonological Theory*. 2nd edition. Chichester, UK: Blackwell Publishing Ltd., 722–750.

Yip, M. 2002. *Tone*. Cambridge, UK: Cambridge University Press.
Yip, M. 2011. Lateral Consonants. In M. van Oostendorp, C. Ewen, E. Hume and K. Rice (eds.) *The Blackwell Companion to Phonology, Vol. I, General Issues and Segmental Phonology*. Chichester, UK: Wiley-Blackwell, 730–756.
Zec, D. 1999. Footed tones and tonal feet: rhythmic constituency in a pitch-accent language. *Phonology* 16, 225–264.
Zsiga, E. C. 2018. The Phonetics and Phonology of English Casual Speech: Learning from L2 Learners. Paper presented at *The Fourth Belgrade International Meeting of English Phoneticians (BIMEP 2018),* English Department, Faculty of Philology, University of Belgrade, 30–31 March 2018.

SUMMARY

CONTEMPORARY PHONOLOGICAL THEORY AND COMMON PRONUNCIATION ISSUES

This paper discusses the role of contemporary phonological theory in teaching English phonetics at university level. In particular, it is shown how, contrary to common popular belief, teaching English phonetics to students of English is not really about teaching them to pronounce English 'properly', whatever the idea of 'proper' English pronunciation might nowadays imply. As university education of future specialists in the English language requires gaining insights into the nature of pronunciation phenomena, it is argued that phonological theory can throw light on some practical pronunciation issues. This is done with reference to theoretical phonological frameworks, such as Laryngeal Phonology, Aperture Theory, Articulatory Theory and Optimality Theory, in order to show how raising phonological awareness about pronunciation phenomena can develop sensitivity to important pronunciation issues.

KEYWORDS: phonological theory, pronunciation issues, teaching English phonetics, phonological awareness.

ARTICLE INFO:
Original research article
Received: June 14, 2018
Revised: February 6, 2019
Accepted: February 9, 2019

UDC: 811.111'342.2
811.134.2'342.2
DOI: https://doi.org/10.18485/philologia.2018.16.16.6

HOW TO HELP LEARNERS TO IMPROVE THEIR ENGLISH PRONUNCIATION. WHAT RIOPLATENSE SPANISH SPEAKING EFL TEACHERS NEED TO KNOW.

MARÍA ALICIA MALDONADO[1]
Santa Fe, Argentina

Ne može se poreći da postoji potreba za predavanjem izgovora u okviru kurseva engleskog kao stranog jezika. S druge strane, dosta se raspravlja o tome da li treba uključiti fonetiku i fonologiju u predavanje engleskog u osnovnom i srednjem obrazovanju. Čini se da pitanje nije da li predavati izgovor, ili šta predavati, već kako predavati. Iz našeg četrdesetogodišnjeg iskustva možemo reći da je neophodno da nastavnici poseduju znanje iz date oblasti ali i da poznaju metodiku, kako bi odabrali prikladan pristup. To podrazumeva znanje o fonološkom sistemu kako sopstvenog L1 (u ovom slučaju rioplatenskog španskog), tako i o fonološkom sistemu engleskog govora koji se predaje, zajedno sa veštinama koje će aktivirati kod učenika "nove načine razmišljanja i konceptualizovanja reči i rečenica stranog jezika" (Fraser 1999: 5). U ovom radu ćemo porediti ova dva fonološka sistema i komentarisati neke od pristupa koji se mogu primeniti kako bi se kod učenika razvili fonološki koncepti koji mogu dovesti do poboljšanja njihovog izgovora engleskog jezika.

Ključne reči: fonetika, fonologija, ELT, španski, engleski.

1. INTRODUCTION

After over 40 years' experience teaching English as a foreign language at all educational levels, out of which 25 were at Teacher Education Programmes and in-service teachers' professional development courses, I can assert that all teachers who aim to help learners to improve their English pronunciation need to have a) expertise in the phonological systems of both English and the native language, b) knowledge

1 Contact information: marilimaldonado@gmail.com

of methodology and teaching strategies and c) a positive predisposition to establish rapport and help students to learn. In this way, learning will become more meaningful, significant and approachable for learners.

In this paper I will address two main issues: I will compare the Rioplatense Spanish and BBC English phonological systems to identify the areas which present most difficulties for Rioplatense Spanish speakers (RSS) to learn English. Besides, I will describe some teaching strategies that might help these students to develop phonological awareness and improve their English pronunciation.

For the sake of this work, when I speak of Spanish, I will be making reference to Rioplatense Spanish (RS). Rioplatense Spanish is the Spanish dialect spoken in the La Plata river basin area in Argentina (Figure 1), namely Buenos Aires, Santa Fe, Entre Ríos provinces and also in Uruguay.

Figure 1. La Plata river basin area in Argentina
(HISPANISMO.ORG)

Due to the fact that it is the dialect used in the media, it is widely spread all over Argentina and Uruguay. Therefore we could assert that it is spoken by most of the inhabitants in both countries. In spite of this, we cannot deny the existence of some regional differences and idiolects. In the analysis that follows, we will consider the standard version of Rioplatense Spanish - the variety foreign students are taught when they want to communicate fluently in Argentina.

BBC English is defined by Prof. John Wells in the *Longman Pronunciation Dictionary* (2008: xvii) as a modernized version of RP. Nevertheless, he points out that "now that BBC admits regional accents among announcers this name has become less appropriate." We understand by BBC English the standard model of correct pronunciation that is taught when speakers of other languages learn English which is not American [we could add Australian, South African, or any other] accent.

RSS teachers who want to teach BBC English to Argentine students need to study the linguistic systems of Spanish and English to detect the areas which students might find most difficult. Even though teachers are not expected to focus on the teaching of

Phonetics and Phonology *per se*, we understand that intelligibility in interaction is only possible if individuals at least approach the target model of pronunciation.

In order to find these areas of difficulty, I will compare Rioplatense Spanish and BBC English. In this paper, I will use terms Rioplatense Spanish/Spanish and BBC English/ English interchangeably, always referring to the same accent.

2. RIOPLATENSE SPANISH AND BBC ENGLISH COMPARED

There are some characteristics of both languages that cause difficulty for RSS students when they are learning English. First and foremost, these are spelling differences. Spanish is characterized by grapheme-phoneme correspondence, which facilitate pronunciation learning. On the other hand, there is no grapheme-phoneme correspondence in English, which poses a great challenge for Rioplatense Spanish Speakers (RSSs) learning English pronunciation.

Since Spanish is a syllable-timed language, RSSs tend to produce full vowels in all syllables and, thus, they do not produce weak forms of items. RSSs need to learn the rhythm of English and the value of the weak vowels /ɪ/, /ʊ/ and /ə/, since they help to give English one of its main characteristics: its stress-timed rhythm.

The flexibility of the Spanish grammar allows Spanish speakers to place the most important piece of information at the end of the utterance. This implies that we usually produce broad focus. On the other hand, English has a stricter grammar system. This makes it necessary for English speakers to mark narrow focus in order to highlight the most relevant piece of information in their utterances.

Both languages share proclaiming and referring tones (Brazil 1994), or falling and rising as these tones are described by Wells (2006), with their generalities as regards meaning. In both cases we associate proclaiming or falling tones with completeness and definiteness and referring or rising tones with incompleteness and evocative value. There are still some differences which bring about misunderstanding when the tones are used wrongly.

2.1 COMPARING RIOPLATENSE SPANISH AND BBC ENGLISH SEGMENTALS

2.1.1 Vowels

2.1.1.1 Monophthongs

Rioplatense Spanish has five monophthongs whose allophones cover a wide area of articulation. This contrasts with the English monophthongs, which are twelve in number (see Figures 2 and 3).

Finch and Ortiz Lira (1982) highlight the striking difference in number between the twelve pure English vowels and the five Spanish ones.

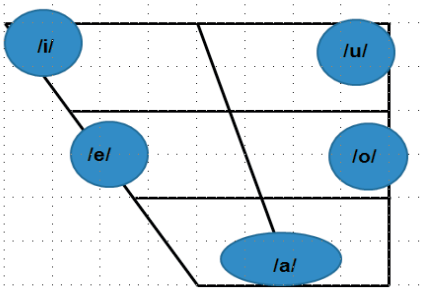


Figure 2. Rioplatense Spanish monophthongs

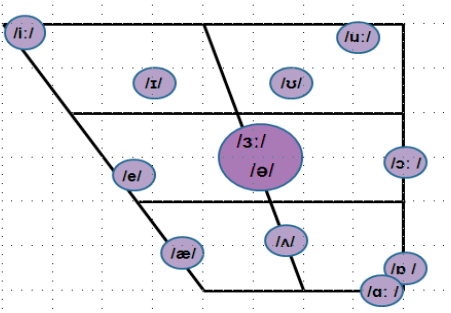


Figure 3. BBC English monophthongs

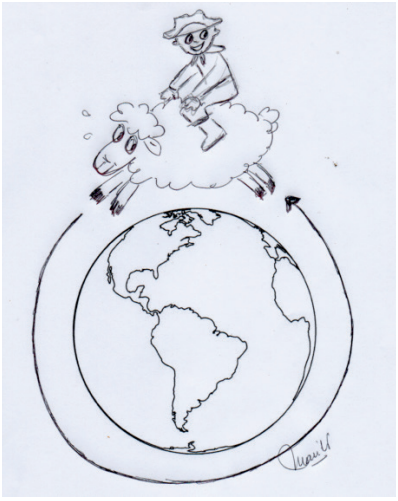
Escudero and Boersma (2004), claim that "the most obvious effect of phonological transfer at the segmental level concerns the difficulty of perceiving the distinction between two sounds in a second language that are not in phonemic contrast in the native language." Observing Figures 2 and 3 it is easy to understand why RSSs find it so difficult to discriminate between English vowels such as /ɪ/ and /iː/; /ʌ/, /æ/ and /ɑː/, or /uː/ and /ʊ/.

Working with examples of minimal pairs in which words share grammatical category is an excellent way to help students to raise awareness of meaningful differences. Showing RSSs how they generate misunderstanding because of mispronunciation of a vowel sound helps them to perceive the sound contrasts. If we work on reactive teaching, i.e. we react to students' production as a native speaker of English would react, RSSs will be able to perceive their mistake more accurately (Kelly 2003; Fraser 1999, 2006, 2008), e.g. a student says *[aɪ ˈkɑːnt liːv wɪðaʊt ˈwɔːtə] and the teacher answers 'Yes, you can!'. We can make the correction more memorable if we add some humour, e.g. when a student says *[aɪm ˈgəʊɪŋ raʊnd ðə ˈwɜːld ɒn ə ˈʃiːp] with the intention of meaning [aɪm ˈgəʊɪŋ raʊnd ðə ˈwɜːld ɒn ə ˈʃɪp], and the teacher answers 'Poor sheep!'. These examples respond to what is usually called a "shocking technique". Students find their teacher's response so unexpected that they have to draw conclusions as regards it and that helps them to become aware of the difference in meaning immediately. This awareness helps learners to improve their production of the differing sounds.

It is still more difficult for RSSs to perceive and produce central vowels since they do not exist in Rioplatense Spanish. Spanish does not have any vowels produced by raising the central part of the tongue. To help students visualize the phonological concept, I suggest we should use the "see-saw" technique (Kenworthy 1987). It consists in producing a cline between two extreme vowel sounds, e.g. move from /iː/ to /ɑː/. While students move their tongue, the teacher makes them stop midway and helps them to perceive the position of their articulator. Immediately afterwards we put a name to the sound which reminds learners of its characteristics. Peter Roach (2005) mentions that the long central vowel is widely known as "the hesitation sound" usually spelt *'er'*. The schwa /ə/, or short central vowel is only produced in a string of speech since it never occurs in stressed position. Because of this reason, and because I adhere

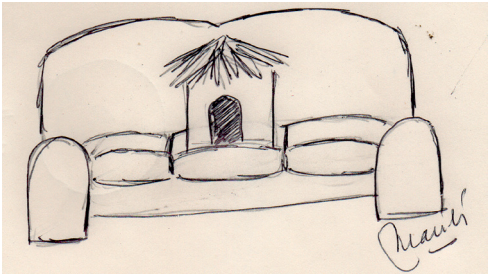
to the wholistic principle presented by Cognitive/Functional Linguistics, I consider it is almost always better to practise pronunciation by producing oral messages. It is in this way that students will become better aware of how their pronunciation mistakes might bring about misunderstanding. Moreover, practising in context, learners will acquire one of the most relevant characteristics of English: its stress-timed rhythm.

Besides the abovementioned techniques, helping students to visualize what they are actually producing instead of what they intend to produce is very helpful as well. Images are a very useful tool to achieve this aim (Figure 4 and Figure 5).



[ˈgəʊ raʊnd ðə ˈwɜːld ɒn a ˈʃiːp]

Figure 4. Visualization of the wrong message: /ʃiːp/ instead of /ʃɪp/



[ðə ˈhʌt ɪz ɒn ðə ˈsəʊfə]

Figure 4. Visualization of the wrong message: /hʌt/ instead of /hæt/

2.1.1.2 Diphthongs

We define diphthongs as vowel sounds consisting of an intentional glide. As a consequence, the first remarkable difference between Rioplatense Spanish and BBC English is the fact that English diphthongs fall into two categories, namely closing and centring, whereas Spanish diphthongs are all closing. Another difference is the fact that English diphthongs have shortened allophones before voiceless consonants or unstressed syllables and Rioplatense Spanish does not have such an allophonic variant.

Both Spanish and English diphthongs are falling because the first element is more prominent than the second one, but Spanish diphthongs are faster and tenser than the English ones. In the production of Spanish diphthongs, the tongue moves towards a closer position reaching the position where the second element is produced. (Figure 6) On the other hand, in English, the tongue moves in the direction of the position where the second sound is produced, but never reaches the point. (Figure 7)

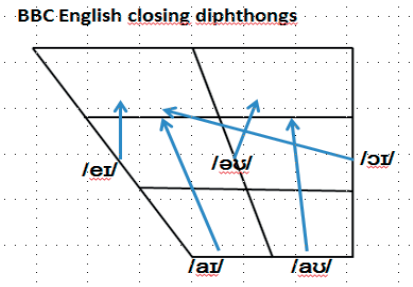


Figure 6. Rioplatense Spanish closing diphthongs

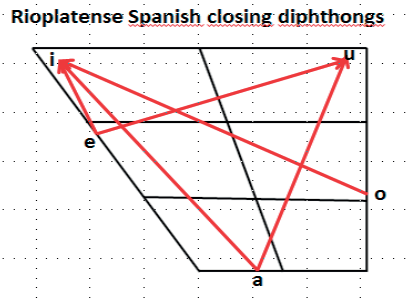


Figure 7. BBC English closing diphthongs

There are no centring diphthongs in Spanish as opposed to English, where there are three gliding towards /ə/, though /ʊə/ is becoming increasingly rare. It is being replaced by the long vowel /ɔː/ as in *poor* /pɔː/. As Collins and Mees (2008) claim "the diphthong /ʊə/ hardly exists".

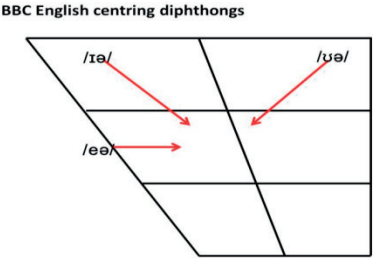


Figure 8. BBC English centring diphthongs

2.1.1.3 Triphthongs

Triphthongs are considered "the most complex English sounds of the vowel type" (Roach 2006). They are the sequence of the five closing diphthongs followed by /ə/. Both Roach (2006) and Collins and Mees (2008) explain that the main difficulty for people learning English lies in the reduction, and even absence of the /ɪ/ and /ʊ/ in the sequence resulting from what Wells (1982) calls 'smoothing' and Finch and Ortíz Lira (1982) describe as "a form of compression technically known as 'levelling'". Spanish speakers tend to produce three-sound sequences called triphthongs which constitute one syllable each. Due to this conceptual difference, RSSs tend to make a very noticeable tongue movement which is perceived by a BBC English native speaker as an intrusive semivowel in the middle of the sequence. To avoid this strong foreign accent effect, RSSs might omit the /ɪ/ in the sequences /aɪ/ and /eɪ/, or let their tongue move slightly in the direction of /ʊ/ in /əʊ/ before /ə/. In this latter case, they should try to avoid rounding their lips too much so as not to pronounce /uː/ or even /w/.

2.1.2 COMPARING CONSONANTS

The consonant system of any language seems easier to teach because we can describe place and manner of articulation. We can even draw the articulators and, nowadays with the help of ICT (Information and Communication Technology), show the movement in animated images. The perception of pronunciation differences is also easier. If we ask RSS students to 'sound English', they will surely produce some characteristic features of English consonants such as aspiration of plosives, darkening of the lateral or production of a post-alveolar approximant instead of a trill. This can be a starting point to teach the English consonant system, but there is much more to be displayed if we intend to teach accurate pronunciation.

If we compare Rioplatense Spanish consonants and the BBC English ones, we will be able to identify the most troublesome areas for RSS to learn. (See Figure 9)

Manner \ Place	bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palato-alveolar	Palatal	Velar	Glottal
Plosive	p b			t d				k g	
	p b		t̪ d̪					k g	
Fricative		f v	θ ð	s z	ɹ	ʃ ʒ			h
		f		s		ʃ		x	
	β		ð					ɣ	
Affricate						tʃ dʒ			
						tʃ			
Nasal	m			n				ŋ	
	m			n			ɲ		
								ŋ	
Flap					ɾ				
Roll									
					r				
Lateral				l					
				l			ʎ		
Approximant	w				ɹ		j	w	
	w						j	w	

Colour	Category
	BBC English phonemes
	Rioplatense Spanish phonemes
	Rioplatense Spanish allophones

Figure 9. English and Spanish consonant systems compared

Observing the chart, we can perceive that there is numerical correspondence between English and Spanish plosives, nasals and laterals. A relevant contrast can be seen between English and Spanish fricatives. We find four fricatives, which are paired in voiced-voiceless phonemes in English, in contrast with four voiceless fricatives in Spanish. (We can add three voiced fricatives which result from allophonic variants of /b/, /d̪/ and /g/, namely, [β], [ð] and [ɣ]). Even though Spanish speakers are not well

aware of the difference between the phonemes and their allophonic variants, they can produce the sounds. Another sound RSSs get to know, even though they do not normally use it, is /θ/. This sound is taught for the sake of developing literacy skills at early school stages. When teachers introduce the letter 'z' they teach students the Peninsular Spanish sound when they say *zapato* (*shoe*), *zanahoria* (*carrot*), *zorro* (*fox*) to make the difference between the letters 'z' and 's'. This will enable informed teachers to help learners to become aware of the characteristics of the phonemes in order to identify them when heard and to produce them in due contexts.

Spanish does not have consonants which can function as syllabic nuclei (i.e. consonants that take syllabic value). As a consequence, students tend to add an extra /ə/-like sound in words such as *bottle* *[bɒtəl] for [ˈbɒtl̩], or *apple* *[æpəl] for [ˈæpl̩]. Only conscious continuous practice can help students to get these allophonic variants right.

Another difference which poses difficulty for RSSs to be accurate is the lack of initial clusters beginning with /s/ + consonant in Spanish. RSS students tend to add an extra /ə/-like sound before the cluster, thus mispronouncing words such as *speak* *[əsˈpiːk] for [spiːk], *stay* *[əsˈteɪ] for [steɪ], etc. In order to put this right, learners are encouraged to either lengthen the initial /s/ (i.e. to produce [sssspiːk] and [ssssteɪ]; or to make use of junction to avoid the addition of the /ə/ (i.e. [ðə ˈstjʊdnt] or [tu ˈspiːk]. Recalling these examples whenever students make a mistake with the clusters will help them to construe their concepts and embody the correct pronunciation.

2.1.2.1 Plosives

English and Spanish share the same number of plosives. And these plosives are paired into voiced and voiceless counterparts. While Spanish voiced plosives are fully voiced in all positions, they are only fully voiced in English when surrounded by voiced sounds.

As regards articulation, bilabial and velar plosives are the same while RS has dental plosives and BBCE comprises alveolar plosives. This is one of the reasons why RSSs articulate /ð/ wrongly, mainly when it is in initial position in the word. On the other hand, RS plosives /b/, /d/, /g/ are realised as the fricatives [β], [ð], [ɣ] respectively when the sounds appear in intervocalic positions, e.g when they say *dado (dice)* [d̪aðo], *barbijo (surgical mask)* [barβixo] or *galgo (greyhound)* [galɣo].

English voiceless plosives have aspirated release when they are in initial stressed position in a syllable. Spanish does not have such aspirated plosives, and we don't find devoiced semivowels following voiceless plosives in such position, either. In order to help RSS learners visualize the production of these allophones Kelly (2003) and Underhill (1994) suggest putting a thin paper in front of our mouths for it to move when we produce the syllables described, or placing a lighted match to see how the flame flickers.

The correct production of aspirated voiceless plosives will help learners of BBC English to make and perceive the distinction between word sequences such as *might earn – my turn / pea stalks – peace talks.*

Final consonant clusters are only found in English. Because of this, RSSs tend not to pronounce the final consonant and produce [restərɒn] instead of [ˈrestərɒnt], or they add a /ɪ/ or /ə/ between the consonants in the cluster. They say *[pɑːkɪd] instead of [pɑːkt].

2.1.2.2 Nasals

The /m/ presents no articulation problem since it exists both in Rioplatense Spanish and BBC English. The special difficulty appears as regards its distribution. There are only a few loanwords in Spanish ending with /m/, so RSSs tend to produce a /n/ instead. We can profit from the grapheme-phoneme correspondence we find in Spanish and ask students to pay attention to spelling to help them to produce the final consonant correctly.

Even though the /m/ only occurs before homorganic consonants in Spanish, RSS learners of English will not find it difficult to pronounce this sound before /f/ and /v/ since both in English and Spanish we have labio-dental assimilation of the nasal sound e.g. *confite (sweet)* [koɱˈfite] and *invitación (invitation)* [iɱvitasiˈon].

The /n/ has the same articulation and distribution in both languages, so it presents no difficulty for RSSs to produce the English sound.

Even though the /ŋ/ is not a Spanish phoneme, it is realised as an allophonic variant of /n/ in words like [ˈtaŋgo] and [ˈtaŋke]. Teachers can use the syllable-timed characteristic of Spanish to ask students to syllabify the word and stop when reaching the end of the first syllable. Students will have the feeling of how the /ŋ/ is produced. Students will have less problems producing this sound in final position, while they will need more practice to produce the sound when followed by vowels.

2.1.2.3 Fricatives

As it was stated before, there is a relevant contrast between the four English fricatives paired in voiced-voiceless phonemes and the four voiceless fricatives in Spanish. Because of this, RSS have special problems when they try to produce voiced English fricatives.

Even though /v/ does not exist in Spanish because it is replaced by [β], the allophonic variant of /b/ which is used indistinctly when reading 'b' and 'v', RSS have the concept of /v/ because the contrast between /b/ and /v/ is emphasized in initial stages of literacy development to teach the spelling of words like '*varón*' (man) and '*barón*' (baron). This helps teachers to exemplify the sound more easily.

A special point is to be made when we teach /ð/. Since in Rioplatense Spanish we produce dental plosives instead of the English alveolar plosives, and [ð] is an allophonic variant of /d̪/ mainly in intervocalic position (e.g. *dado* (dice) [ˈdaðo], *cada* (each) [ˈkaða]) RSSs tend to produce a /d̪/ in initial position instead of the /ð/, and a /ð/ in intervocalic position instead of a /d/ (e.g.*[d̪i ˈəʊðə] instead of [ði ˈəʊdə] for 'the odour'). In order to help students to produce the correct sounds in the correct position, we should try to ask them to produce them in meaningful contexts in an alternated way (e.g. [ði ˈəʊldə ði ˈædʌlt ðə mɔ: ˈmɒdərət])

2.1.2.4 Approximants

Collins and Mees (2008: 273) define approximant as "a manner of articulation produced with the articulators sufficiently apart for there to be no audible friction". They categorise them into central and lateral approximants. All of them only occur before a vowel sound.

While English has a post-alveolar approximant /ɹ/, also called frictionless continuant because the tip of the tongue never touches the roof of the mouth, we find that Spanish has a flap /ɾ/ and a post-alveolar roll or trill in its consonant system. RSS make a distinction between *pero* (but) and *perro* (dog) whereas English speakers don't produce the trill. But this represents difficulty for English speakers to learn Spanish and not the other way around.

The consonants /j/ and /w/ are phonologically like consonants and phonetically like vowels. This is the reason why they used to be defined as semi-vowels. Peter Roach (2005) asserts that the articulation of /j/ is practically the same as that of /iː/ and the articulation of /w/ is very similar to that of /uː/. Nevertheless, he also claims that we use both phonemes as consonants. They always occur before a vowel and they can be devoiced.

While /j/ has no relevant difference between English and Spanish except for the greater muscular tension RSSs produce, we can describe the Spanish /w/ as stressing the velar quality of the sound in contrast with the labial quality emphasis perceived in the English sound. This difference brings about a mistaken production of the English sound by RSSs. They tend to produce a stop or a fricative velar sound before the /w/ in familiar speech e.g. *aguanto (support)* [aˈɣwanto] and thus, they are bound to transfer this habit to English utterances. A good technique to put this right is to ask students to produce a sequence of syllables starting with /w/ and followed by the five Spanish vowels, i.e. /wa/, /we/, /wi/, /wo/, /wu/. When they pronounce this last syllable they actually become aware of the difference between the approximant /w/ and the vowel sound /uː/. To improve accuracy, it is useful to ask students to compete in preparing the longest possible sentence using this approximant. (A group of students of mine created the sentence *Why was Wendy waiting when we walked with wolves on Wednesday?*) They will have fun and they will be practicing both reception and production of the sound. Getting students involved in the visualization of differences, embodiment of new knowledge and emotion commitment, gives them a great chance to succeed in improving accuracy.

As regards the lateral approximant /l/, both English and Spanish share the characteristics of the clear /l/, i.e. the sound produced in initial or medial position. On the other hand, Spanish lacks the strong version of [ɫ] and RSSs understand this allophonic version as 'typical of English'. Due to this perception, they tend to overgeneralize the use of [ɫ] producing it everywhere in utterances for the sake of "sounding more English-like". The accurate differentiation and production of the two versions is quite easy to put right by using utterances which include both, e.g. *Louise loves living all alone in the middle of the field*. Besides highlighting their different phonological characteristics, we can explain phonotactics to aid accurate pronunciation of them.

2.1.3 Intonation

I cannot close this overview without some reference to the relevance intonation has in meaning conveyance. Although it deserves a thorough description in another paper, I will just include a slight reference to its treatment. Tonality, tonicity and tone trace the realization of the different functions of intonation (Wells 2006). It is not what we say but the way in which we say it what makes a difference (Underhill 1994; Kelly 2003). The use of meaningful messages varying dramatically due to diverse intonation helps raising

awareness in students. We resort to critical thinking and guide students into discovery. e.g. when we focus their attention on the difference in meaning between ‖ *A* ˈ*woman* ‖ *without* ˈ*her* ‖ ˈ*man is* ˈ*nothing* ‖ and ‖ *A* ˈ*woman with*ˈ*out her* ˈ*man* ‖ *is* ˈ*nothing* ‖

Another way of helping students is using humour. A good pun (see Fig. 10) can enlighten the learners.

Figure 10. A phonological pun

This joke can aid raise awareness as regards the correct use of word stress. Students will understand the difference between ˈ*baby* ˈ*oil* (oil made out of babies) and ˈ*baby oil* (oil to be used with babies). This joke also gives teachers the possibility of identifying the use of a referring tone anticipating the conclusion of the idea produced with a proclaiming tone. It can also be noted the use of broad focus in the first utterance (the intonation nucleus is produced on the last content word) and narrow focus in the second one (focus on 'baby', which is not the last content word).

[ɪf ˈɒlɪv̥ ɔɪl ɪz̥ meɪd əv ˌɒlɪvz ‖ ˈwɒt ɪz ˄beɪbi ɔɪl meɪd ɒv ‖ ↗əʊ maɪ ˈgɒd]

3. TEACHING PRIORITIES

In order to plan good lessons integrating pronunciation as an essential issue in meaning conveyance, teachers need to be aware of how to grade the demand they pose on learners. Different phoneticians see priorities from different perspectives. Cruttenden (2001) bases the identification of priorities on frequency of occurrence and word categories. He explains that we need to pay special attention to words students will frequently encounter and use. At the same time, he asserts that mistaking the pronunciation of a sound is more relevant when the change brings about misunderstanding because both words (the one produced and the one the speaker intended to produce) share the same category, e.g. *[ʃɪp] .and [ʃiːp] in Figure 4.

Kenworthy (1987) and Collins and Mees (2008) share similar hierarchies of errors. They describe three categories each, as presented in Figure 11.

In both cases they consider that errors that affect intelligibility are to be considered as the most relevant ones. In terms of teaching principles, they have to be taught focally and students should be penalized in case they do not achieve the correct pronunciation.

Secondly, we find errors which do not affect intelligibility but can give rise to irritation or amusement. We can include within this category those errors which give speech a gross foreign accent.

Last, but not least, are those errors which neither affect comprehensibility nor provoke any kind of reaction. They might even pass unnoticed because people who use a particular English accent speak like that, or the feature is "close enough" to the native feature and thus, native speakers can be used to hearing them. In this case, the features are only to be focused upon if the learner is somebody who is being trained to become a teacher of English or plans to work on the basis of his/her language accuracy.

Collins & Mees (2001)	**Kenworthy (1987)**
Category 1: Errors which lead to a breakdown in intelligibility	**High Priority:** Errors which are vital for intelligibility
Category 2: Errors which give rise to irritation or amusement.	**Low Priority:** Errors which often do not affect intelligibility, e.g. sounds which occur relatively rarely, such as /ʒ/.
Category 3: Errors which provoke few such reactions and may even pass unnoticed.	**Optional attention:** Errors which might contribute to a very noticeable foreign accent, but native speakers are used to them either because: 1.They exist in some regional accents or varieties of English. 2.The feature is "close enough" to the native feature.

Figure 11. Hierarchies of error compared

4. CONCLUSION

Quoting Recamán (1979: 66), we could assert that teachers "should possess 'a sound knowledge' of the phonological and phonetic characteristics both of English and Spanish, a pronunciation of the target language close to the model chosen, and the ability to predict errors and use adequate techniques of correction." Helping students to acquire a

good pronunciation, we will be teaching them how to convey meaning correctly because pronunciation is an integral part of meaning, as well as grammar and syntax.

Teachers are not supposed to stop lessons just to practise pronunciation points, but they are expected to plan lessons in which students will have learning experiences which are meaningful and significant. This will help learners to build their English conceptual frame and use the new language with a comprehensible native-like pronunciation.

REFERENCES

Brazil, D. 1994. *Pronunciation for Advanced Learners of English. Teacher's Book*. Cambridge: Cambridge University Press.

Brazil, D. 1994. *Pronunciation for Advanced Learners of English. Student's Book*. Cambridge: Cambridge University Press.

Collins, B. and I. Mees. 2008. *Practical Phonetics and Phonology. A resource book for students*. 2nd edition. Abingdon: Routledge.

Cruttenden, A. 2001. *Gimson's Pronunciation of English*. 6th revised edition. London: Arnold.

Escudero, P. and P. Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26, 551–585.

Finch, D. and H. Ortiz Lira. 1982. *A Course in English Phonetics for Spanish Speakers*. London: Heinemann.

Fraser, H. 1999. *ESL Pronunciation Teaching: Could it be more effective?* [Internet]. Available at: https://www.cdu.edu.au/walking_talking_texts/documents/Pronunciation-paperbyHelenFraser.doc [08.10.2015].

Fraser, H. 2006. *Helping teachers help students with pronunciation: A cognitive approach*. Australia: University of New England. [Internet]. Available at: http://www.ameprc.mq.edu.au/docs/prospect_journal/volume_21_no_1/21_1_5_Fraser.pdf [08.12.2015].

Fraser, H. 2008. *Cognitive Theory as a Tool for Teaching Pronunciation. University of New England*. Australia. [Internet]. Available at: http://www.linse.uni-due.de/laud-downloadliste/articles/cognitive-theory-as-a-tool-for-teaching-pronunciation.html?file=tl_files/laud/A728_Fraser_komplett.pdf [08.08.2015].

hispanismo.org. *Español Rioplatense*. [Internet]. Available at: http://hispanismo.org/hispanoamerica/18089-espanol-rioplatense.html [07.12. 2017].

Kelly, G. 2003. *How to Teach Pronunciation*. Harlow: Longman.

Kenworthy, J. 1988. *Teaching English Pronunciation*. Harlow: Longman.

Recaman, M. 1979. A *comparative study of the phonological structures of English and Spanish*. Unpublished manuscript.

Roach, P. 2005. *English Phonetics and Phonology. A practical course*. Cambridge: Cambridge University Press.

Underhill, A. 1994. *Sound Foundations*. Oxford: Heinemann.

Wells, J. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Wells, J. 2006. *English Intonation: An Introduction*. Cambridge: Cambridge University Press.

SUMMARY

HOW TO HELP LEARNERS TO IMPROVE THEIR ENGLISH PRONUNCIATION: WHAT RIOPLATENSE SPANISH SPEAKING EFL TEACHERS NEED TO KNOW

We cannot deny the need for teaching pronunciation in EFL courses. Still, there is much discussion as regards whether to include phonetics and phonology in the EFL primary and secondary classrooms. The issue seems to be not whether to teach pronunciation, or what to teach, but rather how to teach it.

From my teaching experience of over 40 years, I can claim that teachers need to have the expertise in the subject and the methodology knowledge that enable them to choose adequate strategies. This implies having knowledge of the phonological systems of both their L1 (in this case, Rioplatense Spanish) and the English accent they are expected to teach, together with the skills to activate in learners "new ways of thinking about or conceptualizing words and sentences in the new language" (Fraser 1999: 5).

In this work I will compare both phonological systems and comment on some useful strategies that can be implemented to help students develop phonological concepts that can foster improvement of their English pronunciation.

KEYWORDS: Phonetics, Phonology, ELT, Spanish, English.

ARTICLE INFO:
Original research article
Received: May 10, 2018
Revised: January 31, 2019
Accepted: February 5, 2019

UDC: 811.163.2'342.344
811.111'342.344
DOI: https://doi.org/10.18485/philologia.2018.16.16.7

ENGLISH WEAK FORMS – A CHALLENGE FOR BULGARIAN LEARNERS OF ENGLISH? A PEDAGOGICAL PERSPECTIVE

NIKOLETA STOYKOVA[1]
Sofia University "St. Kliment Ohridski"
Sofia, Bulgaria

Cilj ovog istraživanja jeste da utvrdi da li bugarskim učenicima engleskog usvajanje i tačna produkcija redukovanih formi u engleskim funkcionalnim rečima predstavljaju problem. Takođe se obraća pažnja na to kako bugarski učenici koriste slabe oblike u vezi sa rasporedom naglašenih slogova u toku govora na engleskom u poređenju sa rasporedom slabih oblika i naglašenih slogova u govoru izvornih govornika. Istraživanje se oslanja na empirijske podatke sakupljene tokom časova sa homogenom grupom od 20 bugarskih učenika engleskog koji su studenti prve godine Engleskih i američkih studija na Sofijskom univerzitetu „Sv. Kliment Ohridski". Ispituje se njihova produkcija slabih oblika u poznatom pasusu vezanog teksta koji sadrži dovoljan broj reči koje obično imaju slabe oblike. Svi učesnici su učili engleski kao strani jezik 4 ili više godina. Usvajanje i produkcija engleskih slabih oblika je problematična oblast za izvorne govornike bugarskog i njihova percepcija i produkcija zahtevaju dalje ispitivanje.

Ključne reči: engleski, bugarski, slabi oblici, naglašenost, L2 učenici, produkcija.

1. INTRODUCTION

When learning a foreign language, the majority of students strive to achieve fluency, while a substantial part also aim at native-like proficiency in it. The latter try to adopt as many new items from the vocabulary stock, and grammatical structures from the foreign language, as possible. They try to imitate the way other people talk, behave and act. However, achieving a native-like proficiency in a foreign language is a daunting task, especially because apart from lexis and grammar, there come the phonetics and phonology of the foreign language as well. Actually, phonetics and phonology play quite

1 Contact information: stoykova.n@gmail.com

a significant role in achieving a certain degree of fluency and many aspects of phonetics and phonology have appeared to be quite difficult for students to master (Flege 1995: 266, Best and Tyler 2007: 16). In the specific case of Bulgarian learners of English, we can also talk about the transfer of L1 features to the L2 (Danchev 1988: 96). In my capacity as an assistant lecturer teaching the course of Practical Phonetics to first year students at Sofia University "St. Kliment Ohridski" for the past 3 years now, at the beginning of every year I start the course by asking the students to fill in a short questionnaire. I then analyse it and see what their expectations are, and most importantly, I gather information whether they have a preferred/least liked/unintelligible accent, whether they have previously been taught phonetics or any aspect of phonetics, etc. Some of the answers to the questions gave rise to some ideas for research. I am going to summarize some data from 30 questionnaires which I have taken into account. All of the students are in their first year of studying English philology and are between 19 and 21 years of age. Since some of them had not indicated their gender, and since it is not relevant for drawing conclusions from this exact questionnaire, I will not consider it as an important feature. The more valuable part comes from the answers to some of the questions given by these 30 students. All of the questions were of the open type and the students could write as much as they thought necessary. 60% of them have indicated that they haven't been taught anything about phonetics before, 27% admit they have never used the pronunciation section of a dictionary, and 60% indicate that speaking a foreign language in front of other people, especially natives, makes them uncomfortable and they feel anxious. 53% of the students share that they have never been misunderstood by other people due to their pronunciation, while 20% say *maybe*. Then, the interesting bit comes, where as an answer to the 9th question 100% of the students ascertain that their pronunciation needs to be improved, and as aspects that need to be worked on, they point at individual sounds, specific words, stress, rhythm, and intonation. One of the reasons given was "to avoid sounding too foreign". This makes me draw two conclusions – first, that my students are self-conscious of their pronunciation and they know they need to work on it, and second, that I will have to help them overcome to some extent the fear of speaking English and teach them how to lose their "foreign" accent as much as possible.

English weak forms are a challenging field for research, especially when it comes to their production and perception by non-native learners. According to some of my students weak forms sound "weird" and "indolent", and their users are "sloppy". During my rather short experience of three years of teaching at university, 90% of all my students were surprised to find out that such forms even existed. However, since Ladefoged and Johnson note that:

"There is, of course, nothing slovenly or lazy about using weak forms. [...] Rather than being labeled lazy, it could be described as being more efficient, in that it conveys the same meaning with less effort. Weak forms and assimilations are common in the speech of every sort of speaker in both Britain and America. Foreigners who make insufficient use of them sound stilted." (Ladefoged and Johnson 2011: 111), it seems that weak forms sound unnatural not only to Bulgarian but to other non-native speakers and learners as well.

As Cruttenden (2014: 321) notes, a foreign language is learnt via words in isolation, pronounced in their citation forms, while in real life connected speech, these same words

undergo slight changes that the non-native speakers need to be aware of, incorporate in their own speech styles and resort to using eventually. Giegerich (1992: 249) also states that "Citation forms are a form of speech that is, to say the least, somewhat idealised". In addition to that, Danchev (1988: 102) rightly observes that learning a foreign language without proper study of its phonetic features at the beginning of a language course will lead to more long-lasting pronunciation errors in comparison with a course that covers those features in some detail. Weak forms are an inseparable part of the speech habits of English speakers and as such they have to be acquired by non-native speakers. By acquisition I mean that they have to be able to both perceive and produce them correctly. According to another scholar, Kelly (2002: 75), "it is important that learners are taught the possible forms of these words when they are introduced because they should be given the opportunity of becoming attuned to them" and "native speakers tend not to notice features of connected speech when they are used, but do notice when they are not". When it comes to reasons why it is vital for students to learn how and when weak forms are used, Roach (2009: 89) gives two main ones: the first one is that native English speakers consider an ""all-strong form" pronunciation unnatural and foreign-sounding" and the second is that "speakers who are not familiar with the use of weak forms are likely to have difficulty understanding speakers who do use weak forms; since practically all native speakers of British English use them, learners of the language need to learn about these weak forms to help them to understand what they hear."

The most extensive work on weak forms that I have come across so far belongs to Obendorfer. He presents a list of about 100 words and provides a full definition of weak forms (Obendorfer 1998: 28):

(1) weak forms are paradigmatically non-basic phonological word forms, or parts of word forms.
(2) they represent morphosyntactic words in certain non-prominent contexts.
(3) their phonological shape is semi-reduced, reduced or cliticized.
(4) they are the products of an ultimately idiosyncratic process.

He also divides them into four major classes (Obendorfer 1998: 46): absolute (*than*; *that*), normal (*at*; *can*), occasional (*any*; *so*), and marginal weakeners (*this*; *no*), noting also that "differences in weakening potential are scalar rather than categorical". However, he claims that absolute and normal weakeners are the ones that have high frequency of occurrence in the spoken language (Obendorfer 1998: 72).

2. MATERIALS AND METHODS

20 first-year students in English and American Studies at Sofia University "St. Kliment Ohridski" were recorded for this production task. All of them have just graduated from high school. I have analysed 20 student recordings, 5 belonging to male and 15 to female speakers. All of the participants are native speakers of Bulgarian with 5 or more years of English language learning experience, being at B2/C1 level according to the European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). The students' age is between 19 and 21, the majority of them being 19 and

20 years old. I have not taken into consideration their respective hometowns since this characteristic is not relevant to the current study. The recordings were made in a phonetic classroom with the Audacity software, using the equipment available in the room – a standard headset with a microphone and a computer with a Windows 10 operating system. The produced sound files were saved in *.wav* format.

The subjects were asked to read a short connected speech passage. They were given instructions to read a short text first in order to get better acquainted with it, and then read it aloud in their preferred accent (which was predominantly RP). Each student was given enough time to study the text in advance and then read it out loud. None of the students were told what the connected speech passage was going to test in order to avoid any bias and unnecessary pressure on their side. The subjects had to click on the "record" button to start and on the "stop" button to finish the recording, then save it as a sound file in the *.wav* format with their name as the file name. After clicking on the stop button, they could listen to their own recordings but were not allowed to repeat the recording.

For the production task, which was the basis for the judgement task, I had chosen the very well-known connected speech passage "The North Wind and the Sun". It serves the purpose of this investigation perfectly well, since it contains 19 words which are expected to appear in their weak forms, and these have 47 occurrences in different environments throughout the whole text. The function word that occurs in the largest number of instances is the definite article "the". None of the function words from the passage is expected to be pronounced in its strong, full form. The list of words which are expected to appear in their reduced forms in the text is the following: (in alphabetical order) *a, and, as, at, be, but, could, he, him, his, of, should, than, that, the, to, was, were, who*.

The sole judge of the recorded items was the author of this paper, a native speaker of Bulgarian.

3. ANALYSIS AND RESULTS

The data were analyzed auditorily by counting the wrongly produced instances of the function word items, where the weak forms of English grammatical words were expected to occur. These were then converted to simple percentages. The recordings were checked multiple times in order to detect and note incorrect rendering of weak forms as their strong form counterparts. For reference and guidance, I used the lists of Roach (2009: 90-95), Dimitrova (2003: 59), and Obendorfer (1998: 206–210), as well as a recording made by a native English speaker. Two additional judges were also asked to check the unclear instances of function words in 6 independent cases. These included the pronunciation of "was" by 3 speakers. The whole text of "The North Wind and the Sun" consists of 113 words in total, 47 of which are expected to be uttered in their weak forms. These 47 words constitute 41.6% of all words, and 7 out of these 47 instances (or 15% of them) were pronounced without a mistake by all 20 speakers. 12 instances, or 25.5% of the 47 weak form instances, were pronounced incorrectly by 4 speakers (the majority of the incorrect productions were made by one and the same speaker – 9). The

words that have more than 4 wrong productions represent 46.8% of all 47 expected weak form words, while the total of wrong productions amounts to 85% of all expected weak forms in the text.

The weak forms pronounced incorrectly as strong forms by all 20 speakers were: one instance of *and,* one instance of *he,* and *should*. I have to make a note that *should* is present only once in the text, while *he* is encountered twice, and the occurrences of *and* are four in number. What is interesting to note is that when it comes to the other 3 instances of *and* 17, 19 and 19 out of the 20 speakers produced its strong form, and when it comes to the other instance of *he,* again only one speaker used the weak form – what these show is that the speakers are inconsistent in their usage of weak forms. Since 2 out of the 3 top mispronounced words have other instances throughout the text, we will look at them more closely. Thus, we note inconsistency on the part of the speakers because Speaker 1 produced two reduced and two strong forms for the 4 different instances of *and*. Speakers 3 and 4 similarly produced 1 reduced and 3 strong forms of the same word.

When we take into account the words *he*(x2), *him, his*(x3), we again note inconsistencies in the production. As mentioned above, speakers 3 and 4 have reduced only one of the instances of *his*, while speaker 1 has not reduced the form of *he* only, having produced the rest of the function words in their reduced forms.

To sum up, when we look at the overall performance of the students and their pronunciation of weak forms, we will have to note the following: the speakers with the lowest percentage of incorrect production of weak forms are Speaker 1 (23.4%), followed by Speakers 3 and 4 with 27.7% and 30% respectively. The speakers who have the highest percentage of incorrectly produced weak forms are Speaker 7 (62%) and Speaker 10 (51%). The rest of the subjects (15 students) show results which range between 32–47%.

Table 1. Wrong productions by all speakers

Lastly, I would like to touch upon two function words that are a part of the text – the one that is encountered the most throughout, namely the definite article *the*, and

the other one that has only one instance – the particle *to*. Both words have two weak forms and the use of either depends on the sound that comes next. For *the* there are the variants [ðə] in front of consonants and [ði] in front of vowels. For *to*, we have [tə] in front of consonants and [tu] in front of vowels. Regarding the pronunciation of *to* in "to confess" only 3 out of the 20 speakers produced the correct weak form in front of a consonant. As to the production of the weak forms of *the*, the results are more complex. There are only two instances in the text, where the weak form in front of vowels [ði] has to be used – in "the other" and in "the attempt". In the first instance of the weak form of *the* in the text in front of a vowel, namely in "the other", six speakers wrongly produced the weak form that occurs before consonants. In the second instance – in "the attempt", there were 12 speakers who produced it the wrong way. Four of these 12 speakers produced the weak form of *the* in both "the other" and "the attempt" as [ðə] instead of [ði]. As to the pronunciation of *the* as [ðə] in front of consonants, 10 of its instances were pronounced wrong as [ði] by one and the same speaker, while four other speakers produced only one of its instances incorrectly. This suggests that the speaker who produced the biggest number of these incorrect weak forms is not aware either of the specific requirements for the pronunciation of the weak forms of *the*, or of the existence of weak forms in general.

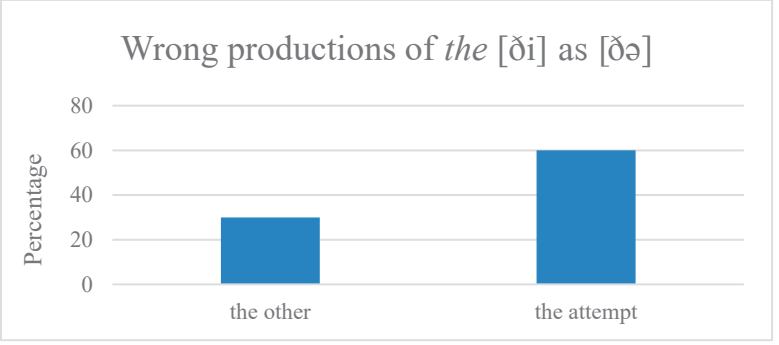


Table 2. Wrong productions of *the* [ði] as [ðə]

4. DISCUSSION AND CONCLUSIONS

The research reported here had as its aim to answer the question whether it is problematic for Bulgarian language learners of English to acquire and produce the reduced forms of English function words correctly.

This study was conducted before all first-year students were explicitly taught a lesson on English weak and strong forms. It was carried out this way, so that their prior knowledge of the topic could be evaluated. Since weak and strong forms are supposed to be taught at secondary level, undergraduates are expected to be aware of their existence and use. What is more, in the questionnaire that I conduct at the beginning of every school year, a number of students admit to having had multiple contacts with native speakers of English, which means that they have been exposed to the native speakers' speech. Taking into consideration all of the above, students, and

the participants in the study in particular, are expected to have a certain amount of knowledge regarding the use of weak and strong forms. However, what also becomes clear from the answers to the questionnaire, is that phonetic instruction at secondary level is cut to a bare minimum. This means that students are not familiar with a number of phonetic aspects, one of which is the issue in question. The results from the current study could serve as a more effective starting point in preparing and teaching lessons on English weak and strong forms to Bulgarian learners of English at tertiary level, thus helping instructors in their practical work.

What became clear to me after conducting the experiment is that the subjects rarely incorporate weak forms in their production task. What they tend to do is use the strong forms instead, or use the wrong variant of the weak form. What is more, for words such as *and* and *that* I heard quite a lot of variation in the production of strong forms, too. Some speakers pronounced *and* as /end/ or /æn/, and a few other speakers produced *that* as /ðet/ or /ðæːt/. This finding goes in tune with the observation that "not only weak forms but strong forms as well vary a great deal in connected speech" made by Makino (2012: 70).

What also became evident for me is that Bulgarian learners incorporate strong forms in places where it is natural for a native speaker to use a weak form instead. Insufficient knowledge, fear not to be misunderstood, and an attempt to blend in and not stand out as a foreigner might be pointed out as some of the reasons for the subjects' use of strong instead of weak forms. In other words, in their wish to drop, they enhance their foreignness instead. Also, the fact that English is defined as a stress-timed language, and Bulgarian as possessing features of both stress- and syllable-timed languages might provide an explanation as to why the distribution of stresses in the flow of the students' Bulgarian English speech is different and sounds chunkier compared to the distribution of stresses and weak forms in the speech of a native English speaker.

This experiment is quite small in scale, and additional follow-up experiments and further in-depth analysis will be needed to verify the present results. However, the current investigation served its purpose in answering the question whether English weak forms present a challenge for Bulgarian learners of the language. The answer is a definite yes, and indicates that their production, as well as perception, will need further research.

REFERENCES

Best, C. T. and M. D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (eds.) *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam: John Benjamins, 13–34.

Cruttenden, A. 2014. *Gimson's Pronunciation of English*. 8th edition. Oxon: Routledge.

Danchev, A. 1988. Kam opisanieto na "balgarskiya anglijski mezhdinen ezik," (Towards the description of the Bulgarian English interlanguage). In S. Peycheva (ed.) *Problemi na ovladyavaneto na chuzhd ezik* (Problems of foreign language acquisition). Sofia: Narodna Prosveta, 93–109.

Dimitrova, S. 2003. *English Pronunciation for Bulgarians*. Sofia: Vezni-4.

Flege, J. E. 1995. Second language speech learning: Theory, findings and problems. In W.

Strange (ed.) *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research*. Timonium: York Press, 233–272.

Giegerich, H. J. 1992. *English phonology: An introduction*. Cambridge: Cambridge University Press.

Kelly, G. 2000. *How to teach pronunciation*. Edinburgh: Pearson Education Limited.

Ladefoged, P. and K. Johnson. 2011. *A Course in Phonetics*. 6th edition. Wadsworth: Cengage Learning.

Makino, T. 2012. Pronunciations in connected speech: a survey of weak forms in a spoken corpus of American English. In T. Paunović & B. Čubrović (eds.) *Exploring English Phonetics*. Newcastle Upon Tyne: Cambridge Scholars Publishing, 57–71.

Obendorfer, R. 1998. *Weak Forms in Present-Day English*. Oslo: Novus Press.

Roach, P. 2009. *English Phonetics and Phonology: A Practical Course*. 4th edition. Cambridge: Cambridge University Press.

SUMMARY

ENGLISH WEAK FORMS – A CHALLENGE FOR BULGARIAN LEARNERS OF ENGLISH? A PEDAGOGICAL PERSPECTIVE

The aim of this study is to determine whether it is problematic for Bulgarian language learners of English to acquire and produce the reduced forms of English function words correctly. What I also note is how Bulgarian learners incorporate weak forms in relation to the distribution of stresses in the flow of their Bulgarian English speech compared to the distribution of stresses and weak forms in the speech of a native English speaker. The study relies on empirical evidence collected in a class environment from a homogenous group of 20 Bulgarian learners of English, who are first-year university students in English and American Studies at Sofia University "St. Kliment Ohridski". I investigate their production of weak forms in a very well-known connected speech diagnostic passage which contains a sufficient number of common weak form words. All of the participants have been studying English as a foreign language for 4 or more years. The acquisition and the production of English weak forms is a problematic area for native speakers of Bulgarian and their perception and production need further investigation.

KEYWORDS: English, Bulgarian, weak forms, stress, L2 learners, production.

ARTICLE INFO:
Original research article
Received: June 14, 2018
Revised: December 1, 2018
Accepted: December 6, 2018

Časopis izlazi jednom godišnje. Rukopise slati na elektronsku adresu:
philologia.journal@gmail.com.
Najlepše Vas molimo da pre nego što predate rukopis pažljivo pročitate Uputstva saradnicima, koja se mogu naći u okviru prezentacije Udruženja na
www.philologia.org.rs
Za sadržaj priloga odgovornost snose sami autori.

The journal comes out annually.
The manuscripts should be submitted electronically at: philologia.journal@gmail.com.
The contributors are requested to refer to the guidelines on the Philologia's website:
www.philologia.org.rs (Submission Guidelines).
The authors bear full responsibility for the contents of their papers.