UDC 811.111'342-057.875(497.11)
Danica Jerotijević Tišma
University of Kragujevac
Faculty of Philology and Arts

ON THE PERCEPTION AND PRODUCTION OF ALLOPHONES BY SERBIAN EFL LEARNERS

Abstract: Well-established research within the framework of the *Speech Learning Model* (Flege 1995) hypothesized that learners might form a new category in the interlanguage phonological system provided that they perceived target contrasts as phonetically distinct from mother tongue sounds. With this in mind, the acquisition of L2 allophones is possible if the perceived distance is sufficient to avoid equivalence classification. The same model likewise presumed a correspondence between the levels of perception and production.
Bearing all this in mind, the present study focuses on the examination of perception and production of English allophones by Serbian EFL learners at the tertiary level of education, i.e. on the effect of phonetic context on the perception and production of target consonants, since allophones of consonants were the only allophones actually taken into consideration. The empirical research consisted of perception and production testing by means of identification tasks and sentence reading which was supposed to simulate spontaneous speech. The obtained results in the perception testing were analyzed in order to obtain the level of accuracy, while the production corpus was examined in terms of the actual production of allophonic variation, i.e. of the practical realization of English allophones in Serbian EFL learners' speech. A total of thirty-five second-year English-major students participated in the study.
The results point to the discrepancy between perception and production, revealing a higher level of accuracy in perception for some allophones, with the actual realizations of allophones in speech being appreciably scarcer. The differences between perception and production may indicate the process of acquisition in progress, but may also challenge presupposed misconceptions on the causality of the two. The concluding remarks of the paper underscore important pedagogical implications for teaching the pronunciation of allophones in an EFL/ESL context.

Key words: allophones, SLA, interlanguage, Serbian EFL learners, phonetic context

1. Introduction

To make communication clear and understandable, speakers generate sounds through motor commands that differ across phonemes, and listeners identify them as meaningful units, i.e. vowels and consonants of a particular language (MacNeilage 2011). The production of phonemes requires a precise control of vocal tract movements and several productions of one and the same phoneme may result in considerable variability due to, on one hand, the variability of the biomechanics of the vocal tract itself, and, on the other hand, the effects of the phonetic context, i.e. surrounding sounds.

When learning a second language, students need to learn not only how sounds make patterns, but also how contextual factors determine changes to their underlying forms. Contextual variability is exactly what shapes allophonic variants of phonemes (Ashby 2011: 12), hence, learners must decide on the similarity and dissimilarity of the segments via specialized cognitive and computational processes. Unlike first language acquisition in which these computational processes happen implicitly within the first year of life, the situation in SLA is more complex and time-consuming, often resulting in a lack of native-like competence, especially for adult learners who are perceptually accustomed to the sound patterns of their native language (Werker, Curtin 2005). Even though researchers have drawn attention to the necessity of instructing L2 learners in the field of subphonemic and allophonic variation (Prator 1971: 71; Standwell 1973: 119), compared to the research in the acquisition of phoneme categories, the research in the SLA of allophones still remains slightly neglected. The reason behind such a state of affairs may be the fact that relevant literature offers fairly little regarding the ways in which non-contrastive, allophonic variants are perceived, whereas the amount of information related to the perception of phonemic contrasts seems everything but scarce (Chappell 2017: 17). The issue of whether allophonic variation should be introduced to everyday ESL/EFL classes is rather disputable, since a teacher needs to determine how deep into phonetic details of sound articulation he/she wants to delve and present to students so as not to overwhelm them with impractical and confusable articulatory details (Collins, Mees 2013). Taking the Critical Period Hypothesis into consideration and having in mind that its main postulate underlines that acquiring a native-like production after puberty is impossible (Young-Scholten 1992: 201), teaching tiny phonetic details and allophonic variations seems a waste of time, when one cannot even discriminate between phonemic categories. Furthermore, one might think that attention paid to intricate details of sounds could perfect gifted students' pronunciation, placing those less gifted into a disfavored position. The question likewise arises pertaining to which allophones should be given a priority in teaching, since some of them are more conspicuous, hence more likely to cause comprehension problems when pronounced incorrectly.

The present paper thus aims at potentially filling the gap in the research related to the acquisition of allophonic variants of L2 consonants, especially in the Serbian EFL teaching and scientific context. To the best of our knowledge, scientific studies related to the acquisition of L2 allophones, especially empirically based ones, are almost non-existent. Having the previously mentioned in mind, the current paper focuses on the perception and production of contextually determined variants in the interlanguage system of Serbian and English, underscoring the most prominent allophonic variants of Contemporary English.

2. Phonetic Context and the Perception and Production of Consonants

In various phonetic environments, phonemes are realized differently and their mutual effects are intertwined.

Speech perception research has shown that the perception of speech sounds is dependent on the sounds surrounding them (Sitek 2011: 77). Identification and discrimination accuracy likewise depend on phonetic context, hence, for example, the /r/-/l/ contrast will be perceived more easily at the ends of words than in initial consonant clusters by Japanese learners (Logan et al. 1991: 883). Examinations of Japanese and English vowel systems, disregarding phonetic context, will be insufficient to explain why assimilation occurs (Strange et al. 2001: 1692).

Clear acoustic features of a sound may change depending on phonetic context, hence, the contrast between /s/ and /ʃ/ may be reduced before a rounded vowel (Mullennix et al. 1989: 368). The effect of the phonetic environment on the movements of vowel formants was frequently investigated, and the research has shown that great transitions of formants (especially the ones belonging to a rounded vowel) happen particularly in alveolar environments (Hillenbrand et al. 2001: 748).

Phonemic and phonetic contrasts are acquired faster in initial positions than the ones in final positions (Eckman, Iverson 1994: 261), which means that certain positions yield less favourable results regarding acquisition than others. However, Japanese EFL learners perceive liquids in final positions better than in initial positions, according to one particular study by Strange (1992: 197-200). It is believed that the overall acquisition of target sounds is at a higher level if students learn sounds in as many different environments as possible.

VOT duration of /t/ will be longer in initial than medial positions (Pierrehumbert, Talkin 1992: 95). English consonants in general show greater articuatory force in initial positions (Fougeron, Keating 1997: 3729), thus we might expect stronger articulation in accented sylables, or in initial positions, than medial ones.

Speaking of the perception of L2 allophones, the presence of similar allophonic categories in listeners' L1 helps perception significantly. Korean has an intervocalic lateral geminate sound, which does not exist in Japanese, which seems to help them map English /l/ and /r/ onto the singleton and geminate allophones of the L1 liquid (Ingram, Park 1998).

3. Acquisition of Allophones in the Framework of the *Speech Learning Model* (Flege 1995) and the *Perceptual Assimilation Model* (Best 1995; Best et al. 2001)

Transfer of phonological features from a native language into an L2 has been recognized as one of the most influential factors in determining target sound production ever since the beginning of SLA research (Trubetzkoy 1939). However, the phenomenon of transfer has since been regarded from various perspectives.

One of the theoretical models seeking to explain production errors and foreign accent is Flege's *Speech Learning Model* or SLM (Flege 1995), stating that L2 learners will have more difficulty in acquiring an L2 sound that is similar to a sound of their native language than a sound that is new and different. Equivalence Classification impedes the formation of a category for an L2 phoneme when it has a similar counterpart

in the L1. The L2 learner perceives the sounds as equivalents which results in the substitution of the L2 sound with the L1 sound. An especially important postulate of SLM for our present research is its first hypothesis which premises that the relation between L1 and L2 sounds does not necessarily occur at an abstract phonemic level, but rather at a position-sensitive allophonic level. L2 learners perceive L2 sounds as allophones and later match them to L1 sounds. This might suggest that L2 learners can detect the phonetic difference between two allophones of a phoneme in L2 and develop a new category.

The *Perceptual Assimilation Model* or PAM (Best 1995; Best et al. 2001) assumes that people perceive sounds based on articulatory gestures, following the concept of direct realism. Even though all people have the same vocal apparatus, they do not employ it in the same way for the production of sounds in different languages. Depending on how L2 learners perceive similarities or differences between the sounds of L1 and L2, L2 sounds will be assimilated to native phonetic categories in various ways.

Even though literature often underscores PAM's treatment of phonemic categories and the difficulties of acquisition at the segmental level, we cannot neglect the fact that Best and her colleagues addressed the issue of allophonic acquisition specifically in one of their later papers, of course relying on underlying ideas of perceptual assimilation (Whalen et al. 1997: 504, 525). The hypotheses of the model in question can be extended to allophones, and contrasting them to native phonetic categories will determine the level of difficulty in their discrimination. If allophones are treated as members of a single phonological category, one might expect single-category assimilation and poor discrimination success. However, if inappropriate and appropriate context distinction is considered, and contextually appropriate allophones are contrasted to contextually inappropriate ones, we would be dealing with a category-goodness assimilation, in which case the inappropriate context allophone would be a less good fit, resulting in better discrimination.

One of the rare studies of L2 allophone acquisition, investigating stop-approximant alternation of Spanish phonemes /b d g/, came to the conclusion that learners' representations of phonetic categories become more precise and detailed as their experience in L2 grows (Shea, Curtin 2011). More experienced learners seem to employ several factors in production, such as the position of a sound in a word, lexical stress, the specific nature of a consonant being pronounced, etc.

4. Allophones in English

Allophonic variation is predictable since allophones are in complementary distribution (Gussenhoven, Jacobs 2011: 63). The task of foreign language learners is to detect the differences and similarities among contextually dependent variants (Werker, Curtin 2005; Peperkamp et al. 2006).

Picard (2007: 337) speaks about the most frequent allophones found in the American English variety, including t-tapping and aspiration. In this section of the paper we will briefly list some of the most common allophonic variations in English,

including both British and American English varieties, since the two are reportedly most widespread among Serbian EFL learners, especially the latter one. We will present only the ones that were included in the empirical segment of the paper, since other allophonic variations exceed the scope of the current paper.

Voiceless plosives /p/, /t/, /k/ are aspirated in a pre-vocalic accented initial position (Ashby 2011: 125), e.g. [pʰi:z], whereas in a word-final position, aspiration is optional. Aspiration is perceptually detectable by longer VOT (*Voice Onset Time*), a phonetic parameter marking the period from the burst of the plosive until the beginning of the periodic vibrations of the following vowel. Hence, in English VOT is generally longer than in Serbian.

Especially in American English, alveolar plosives, or the consonant cluster nasal + plosive, are pronounced using the tapped pronunciation especially intervocalically, e.g. *cider* [ˈsaɪɾɚ] (Odgen 2009: 114).

Two plosives found next to each other result in the first one being unreleased, e.g. *stabbed* [stæbˀd].

The pronunciation of /d/ and /t/ will be retracted before the retroflex pronunciation of the rhotic approximant /r/, e.g. *try* [t̪̊ɾaɪ]. Older speakers still resort to pronouncing the alveolar variant in these environments.

Nasals will be syllabic if preceded by an obstruent, e.g. *prism* [ˈprɪzm̩].

The usual palatalized pronunciation of the lateral consonant („clear" [l]) precedes a vowel, while in pre-consonantal context, the pronunciation of the lateral will be velarized, resulting in the so-called *dark* [ɬ]. Cockney, for instance, exhibits such a high level of velarization that the lateral turns into the low-to-mid back monophthong, e.g. *milk* [mɪɬk] becomes [mɪɒk] (*l-vocalization*) (Wells 1982: 302).

Glottal stops, especially prominent in British English accents and dialects, can either emphasize the articulation of voiceless plosives, or substitute them in intervocalic positions, especially in dialects. Despite widespread misconception, glottal stops can be found in American varieties as well, especially word-finally.

5. Methodology

5.1 Aim of the Study

With the importance of acquiring L2 allophones for achieving native-like pronunciation in mind, the aim of the present study was to investigate the Serbian EFL learners' acquisition of L2 allophonic variation, focusing precisely on the perception and production of English allophones. Furthermore, the goal was to compare levels of perception and production, and possibly gain insight into the nature of their interrelationship.

5.2. Research Questions

The paper was based upon the following research questions:

- Are Serbian EFL learners at the tertiary level of education able to correctly perceive English allophones? How accurate is their perception?

- Are Serbian EFL learners at the tertiary level of education able to produce native-like English allophones? How accurate/native-like is their production?
- What is the relationship between perception and production of allophones?

5.3. Participants

A total of 35 second-year English major students at the Faculty of Philology and Arts, University of Kragujevac participated in the study (academic year 2017/2018, mean age 20.34, 8 male, 27 female). At the time of testing the students were attending the *English Phonology* course, which devotes several introductory units to theoretical considerations related to the perception and production of allophones, as well as to practical exercises where students have time to adopt the presented concepts through various activities. The participants were chosen based on their performance on the mid-term test which consisted of questions related to allophones and features of connected speech with a maximum score of 40 points (20 of which were for allophones only). Hence, the sampled participants performed with 75% or more on the part of the test related to allophones.

5.4. Instruments and Procedure

The perception testing instrument was an identification task that contained 21 sentences with one target token, i.e. English allophone. There were three sentences per allophone, which means that the number of the tested allophonic variations was seven. The allophones were chosen based on their prominence, frequency and analysis accessibility: aspiration, 'l' velarization, t-tapping, retraction, lack of explosion, syllabicity, glottal stopping. The task for the students was to listen to pre-recorded sentences (pronounced by an American English speaker, since the sample of our participants opted for using the American variety). The examples of glottal stops were thus given in pre-nasal and post-vocalic word-final contexts only. The sentences containing target tokens were underlined so as to avoid confusion with other possible allophones in the sentence.

The same allophones were tested in production, except that this time, the students' task was to read the target sentences (differing from the ones in the perception task, of course) in a careful and natural way, three times in a row. Again, each allophonic variation appeared in three different sentences, amounting to 21 target sentences overall.

5.5. Procedure and Data Processing

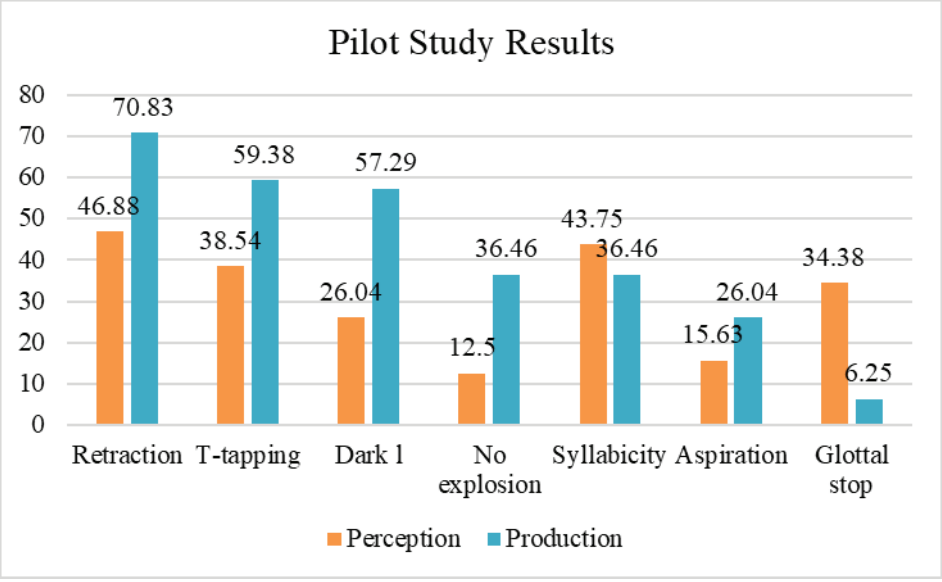
The testing of perception and production was performed during the summer semester 2017/2018 academic year, more precisely, on several separate occasions in May 2018. During the perception testing, the students listened to the pre-recorded sentences and circled their answers on answer sheets. Regarding production testing, students were recorded pronouncing the sentences with target tokens as naturally as possible, three times in a row. The recording was done using an Olympus Digital Voice Recorder VN-8600PC with 16-bit conversion and 44.1 kHz sampling. No indicators were provided related to the places one would expect to find an allophonic variant.

After the testing was done, the perception test was analyzed by percentage counts and correct responses, while students' production was analyzed acoustically and auditively by the author of the paper using Praat, version 6.02.3 (Boersma, Weenink 2016). Perception and production scores comparison was performed using a student's t-test in SPSS, version 23.0.

5.6 A Pilot Study

Before conducting the actual empirical testing for the needs of the present paper, we performed a small-scale pilot test with 32 first-year English major students at the Faculty of Philology and Arts, University of Kragujevac (academic year 2017/2018; 12 male, 20 female, mean age 19.56). The participants underwent the same testing procedure with the identical perception and production tasks. They were formally and practically familiarized with the necessary terms and concepts related to the testing content during the college courses in phonetics and phonology. The results of the perception and production tasks of the pilot study are presented in Graph 1.

Graph 1. Results of the Perception and Production in the Pilot Study



Having in mind that the first-year students had no prior knowledge of allophones, since the topic is introduced during their second year, we excluded the results from the overall discussion of the paper and resorted to testing the second-year students only. We believed that the results were not the display of the actual state of affairs and were unfair to students, since they relied merely on what they had received from inside or outside primary and secondary schools, without being introduced to theoretical, formal, and practical instruction related to allophones. Hence the results were expectedly low for some allophonic variants both for perception and

production, especially regarding aspiration, lack of explosion and l-velarization. It is interesting to note the discrepant results in perception and production with the production scores being higher for the majority of allophones, which goes against the assumption that perception is a necessary prerequisite for successful production. We may ascribe such a state of affairs in the pilot study to the unfamiliarity of students with the perception testing task, since the students were tested for identification for the first time. Furthermore, the discrepancy may have resulted from students' lack of formal knowledge in allophony, so they knew what to produce based on their experience with the target language, but not what to perceive. Nevertheless, the pilot study enabled us to gain insight into the applicability of the proposed instruments and testing materials.

6. Results and Discussion

Table 1 presents the results of the perception testing.

Table 1. Perception Task Testing Results

Allophonic Variation	**Perception Accuracy**	
Retraction		
/d/ → [d̠]/ #_ ɹ	77.14%	N=35
/t/ → [t̠]/ #_ ɹ	84.29%	N=70
T-tapping		
/t/ → [ɾ] / V_V	74.29%	N=35
/ V_n	54.28%	N=35
/V_m	62.86%	N=35
L-velarization		
/l/ → [ɬ] / V_#	60%	N=70
/V_C	45.71%	N=35
Lack of Explosion		
/b/→ [bˀ] / V_d	31.43%	N=35
/d/→ [dˀ] / V_b	28.57%	N=35
/g/→ [gˀ] / V_d	40%	N=35

Syllabicity		
/m/→ [m̩] / z_#	34.28%	N=35
/n/→ [n̩] / d, z_#	68.57%	N=70
Aspiration		
/p/ → [p^h]/ '#_V	77.14%	N=35
/t/ → [t^h]/ '#_V	68.57%	N=35
/k/ → [k^h]/ '#_V	62.86%	N=35
Glottal Stop		
/t/ → [ʔ] / V_n	57.14%	N=70
/ V_#	48.57%	N=35
Within Subject T-test	t=10.439 p=0.001 SD=15.03	

The highest level of perception is noted with allophonic variations of retraction, t-tapping and aspiration, which is expected for the first two allophonic variations and may be the result of a frequent appearance of these allophones in the input. The relatively high score for aspiration is slightly surprising since it probably resulted from formal instruction, i.e. this is the first time students became aware of the existence of this allophonic variation. However, they may have been perceptually but unconsciously accustomed to hearing aspirated sequences, hence their perceptual apparatus recognized phonetic features such as a longer VOT. Retraction shows a slightly higher score for the voiceless alveolar plosive, while t-tapping shows the highest level of identification in the intervocalic position, which again may be due to the students' linguistic experience with American English. Aspiration displays the highest score for the bilabial voiceless plosive.

Surprisingly, l-velarization displays an accuracy score around 60%, which may be lower than expected since the variation is fairly prominent and there is even a similar alternation in the participants' mother tongue (known as "l" to "o" alternation). However, the similarities in allophonic variations may be exactly what hinders accurate perception, according to SLM's suggestions (Flege 1995). Having in mind that glottal stopping requires a certain amount of articulatory effort and force, it was relatively easily identifiable for the participants in our study. The lack of explosion, or unreleased plosives were the most difficult to perceive, possibly because of their lack of audible release, causing perceptual confusion. Interestingly, syllabicity shows quite a discrepant perceptual accuracy depending on the context. Namely, students recognized the alveolar nasal in syllabic environments slightly better than the bilabial nasal, which may be ascribed to lexical frequency. Overall,

the current sample of participants displayed considerably higher levels of perception of allophonic variations compared to the pilot study group, which points to the importance of formally introducing the issue of allophony in the EFL curriculum.

Judging by the results from Table 1, we could say that a particular phonetic environment either facilitated perception or made it more difficult. Within-subject statistical testing (t=10.439 p=0.001) showed that there was a statistically significant difference in the perception of various examples of allophonic variation, which means that not all allophonic variations are perceived equally well by Serbian EFL learners, even though an equal amount of attention was paid to each during formal instruction of *English Phonology*. This points to the intricate inteconnections of various factors contributing to the level of perception of allophones in the interlanguage.

The results of the production testing are presented in Table 2.

Table 2. Production Task Testing Results

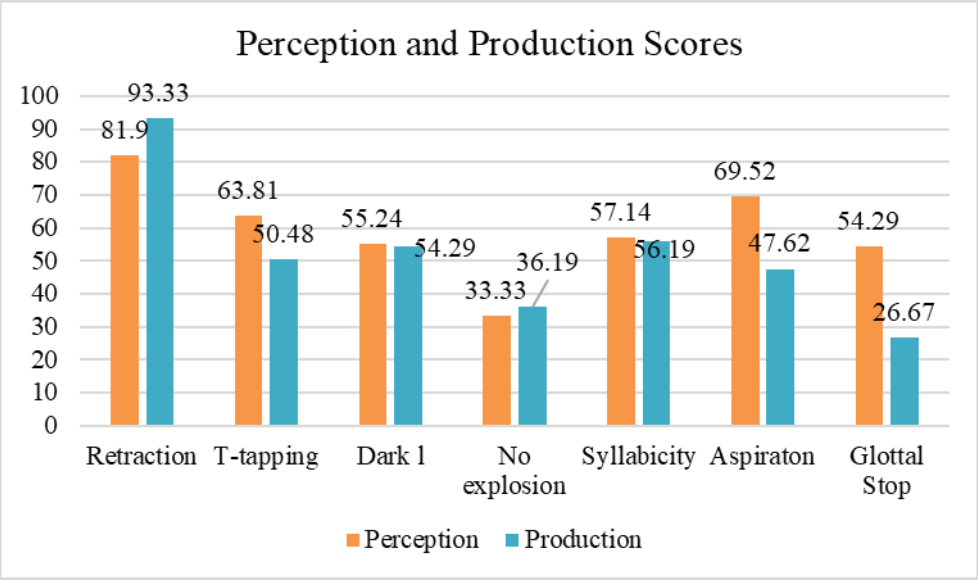
Allophonic Variation	**Production Accuracy**	
Retraction		
/d/ → [d̪]/ #_ ɹ	94.29%	N=35
/t/ → [t̪]/ #_ ɹ	92.86%	N=70
T-tapping		
/t/ → [ɾ] / V_V	54.29%	N=35
/ V_n	42.86%	N=35
/V_m	54.29%	N=35
L-velarization		
/l/ → [ɫ] / V_#	55.71%	N=70
/V_C	51.43%	N=35
Lack of Explosion		
/b/→ [bˀ] / V_d	37.14%	N=35
/d/→ [dˀ] / V_b	28.57%	N=35
/g/→ [gˀ] / V_b	42.86%	N=35
Syllabicity		
/m/→ [m̩] / t, z_#	62.86%	N=70
/n/→ [n̩] / d _#	42.86%	N=35

Aspiration		
/p/ → [p^h]/ '#_V	48.57%	N=35
/t/ → [t^h]/ '#_V	42.86%	N=35
/k/ → [k^h]/ '#_V	51.43%	N=35
Glottal Stop		
/t/ → [?] / V_n	25.71%	N=70
/ V_#	28.57%	N=35
Within Subject T-test	t=6.568 p=0.001 SD=20.99	

The results of production testing singled out retraction as the allophonic variation with the highest frequency of occurence produced in a native-like manner. The scores are similar for both the voiced and voiceless alveolar plosives. This means that Serbian EFL students produce retracted variants almost every time, with very few students resorting to alveolar articulation. This is understandable having in mind that the relevant literature underscores an ongoing sound change, especially in American English, with the widespread affrication of plosives in this particular environment (Magloughlin 2018). The following allophonic variations exhibit a relatively similar level of production accuracy: syllabicity, aspiration, t-tapping and l-velarization (around 50%). Syllabicity shows higher accuracy within the bilabial context, which opposes the results in perception. L-velarization is slightly higher in word-final positions, while t-tapping shows the lowest score in a pre-alveolar nasal environment. Unlike perception, the highest scoring variant for aspiration appears to be the velar plosive. Lack of explosion is one of the lowest scoring allophonic variations especially for the voiced alveolar plosive, while the least produced allophonic variant is definitely the glottal stop, which appears in less than one-third of the cases. Again, within-subject statistical testing showed that there was a statistically significant difference in the production of different allophonic variations, which means that students do not pronounce all of the allophonic variants equally, i.e. some of them are used more frequently than others. Various intertwined linguistic and extralinguistic factors contribute to such a state of affairs, yet we will examine a seemingly obvious one, i.e. the relationship between perception and production.

For the sake of clarity and better overview, we compare the results of perception and production by visually illustrating the scores in Graph 2.

Graph 2. Perception and Production Tasks Comparison



Unlike the pilot study, which showed an appreciably higher degree of discrepancy of scores for perception and production, with the sampled group of participants in the present research we notice that the relationship is slightly better-leveled yet not without its own peculiarities. Namely, for t-tapping, l-velarization, syllabicity, aspiration and glottal stopping, we notice that perception displayed higher scores than production. For aspiration, and especially for glottal occlusion, we notice that production scores are much lower than perception scores. This means that students perceived the given allophonic variants better than they produced them, which is expected if we assume the linear relationship between perception and production, i.e. that perception precedes production, yet longitudinal research is necessary for confirming these claims. However, if students needed to perceive allophonic variants correctly in order to be able to produce them similar to native speakers, the results of the present study could mean that the students are currently in an ongoing process of acquisition which is why their production scores are lower than perception. The previous assumption, though alluring, does not seem to apply to the remaining examples of allophonic variation. In the case of retraction and lack of audible release, the perception scores are lower than the production scores. The reason behind this may lie in either the testing design or in the complex nature of the interlanguage phonological and phonetic systems. More precisely, the students may be able to produce the sounds that they receive from the input because their articulatory organs rely on imitative motor movements, yet they are not able to perceive, i.e. correctly recognize the phonetic features in a given context. For these two specific cases production seems to precede perception, yet again we underscore the need for a more detailed longitudinal approach to the issue. Between-subject

t-test results did not show a statistically significant difference between the results of perception and production (t=0.739 p=0.474).

Nevertheless, the discrepancies in perception and production emphasize the complexity of the interlanguage system, and may point to an even more intricate and demanding situation with the acquisition of allophones than is the case with phonemes. When acquiring phonemes, students involve complex processes of identification and discrimination of phonemic and subphonemic elements, as well as native language attunement, yet with the acquisition of allophones the situation may only be more complex, because it needs to incorporate fine phonetic details, information on the effect of the phonetic environment, as well as phonotactic contraints. Following the hypotheses of SLM (Flege 1995) and PAM (Best 1995), students need to discriminate between native and foreign language phonetic features, but if they ascribe them to a single category, then their overall perception will be hindered. More experimental research is necessary to see how the phenomenon of equivalence classification works at the finely grained phonetic level.

7. Conclusion

The present paper represents an empirical investigation of the perception and production of allophonic variation by Serbian EFL learners, i.e. second-year English-major students at the Faculty of Philology and Arts, University of Kragujevac. Taking the lack of research dealing with allophonic acquisition in the Serbian scientific context into consideration, as well as the impossibilty of attaining target-like pronunciation, the present study focused on testing the accuracy of allophone identification and production. Aware that tackling in such fine detail L2 pronunciation goes directly against the notions of intelligibility and comprehensibility as aims of SLA, we must underscore that the sample of chosen participants was represented by philologists, future teachers, whose pronunciation aims, we believe, should be set slightly higher than merely achieving inteligibility of the output.

The results of perception showed that Serbian EFL students identified retraction, t-tapping and aspiration most successfully, with the lack of explosion posing the greatest difficulty. We ascribed such a state of affairs to the frequency of input and linguistic experience, as well as to the specificity of the allophone with no audible release making identification more demanding. The results of production demonstrated that retraction was produced most frequently with syllabacity, aspiration, t-tapping and l-velarization, displaying around 50% of native-like production. The glottal stop appears to be scarce in the corpus, probably again due to the linguistic experience of our participants. Specific phonetic environments proved important since they either hindered or facilitated both the identification and production of allophones. The relationship between perception and production raised important issues to be potentially answered in future studies. Namely, for the majority of allophonic variations, perception scores were higher, which pointed to

the fact that students were in the process of acquisition, with production waiting to approach the accuracy level of perception.

However, retraction and lack of explosion showed opposing results, with production scores higher than perception scores. This may have resulted from a methodological issue, with perception tasks being more demanding, and less familiar than production tasks. Nevertheless, we allow for the interpretation by which the discrepant results point to the complexity of SLA of subphonemic elements, involving an intertwined network of a multitude of linguistic and extralinguistic factors. Longitudinal studies dealing with perception and production are the only reliable ways to investigate the nature of their interrelationship.

That students generally display a relatively low level of the production of allophones may be explained by resorting to the postulates of the Speech Learning Model (Flege 1995) and the Perceptual Assimilation Model (Best 1995). More precisely, impeded by the mechanism known as equivalence classification and perceptual assimilation, students tend to attune novel phonetic features to already familiar mother tongue features, and, depending on how well they manage to discriminate between them, their perception and production will approach native speakers' ability. The aforementioned models offer explanations for L2 phonemic and subphonemic acquisition, yet it seems that the research dealing with the acqisition of allophones needs to be approached from additional perspectives, such as through investigating complex cognitive and perceptual mechanisms. Hence, future studies may focus on investigating the perception and production of allophones by using both formal and spontaneous speech testing tools, as well as more detailed perceptual tests.

Nevertheless, the study pointed to the importance and complexity of L2 allophone acquisition, underscoring the need for introducing formal instruction and authentic input even earlier than tertiary education. Careful planning, theoretical and practical implementation of the materials on allophonic variation may positively affect both perception and production in the long run, and prepare students for future pronunciation challenges.

References

Ashby, P. 2011. *Understanding Phonetics (Understanding Language)*. New York: Routledge.

Best, C. T., McRoberts, G. W., Goodell, E. 2001. Discrimination of Non-Native Consonant Contrasts Varying in Perceptual Assimilation to the Listener's Native Phonological System. *Journal of the Acoustical Society of America*, 109 (2), 775-794.

Best, C.T. 1995. A direct realist view of cross-language speech perception. In: Strange, W. (ed.), *Speech Perception and Linguistic Experience. Theoretical and Methodological Issues.* Baltimore: York Press, 171–203.

Boersma, B., Weenink, D. 2018. Praat: Doing phonetics by computer (Version 6.04.0) [Computer program]. http://www.praat.org/ [2018, May 11]

Chappell, W. 2017. Costa Rican Spanish speakers' phonetic discrimination. *Estudios de Fonetica Experimental*, XXVI, 13-62.

Collins, B., Mees, I.M. 2013. *Practical phonetics and phonology.* London: Routledge.

Eckman, F. R., Iverson, G. K. 1994. Pronunciation difficulties in ESL: Coda consonants in English interlanguage. In H. Yavas (Ed.), *First and Second Language Phonology* (pp. 251-265). California: Singular Publishing Group, Inc.

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* (pp. 233–277). Baltimore, MD: York Press.

Fougeron, C., Keating, P. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 101, 3728-3740.

Gussenhoven, C. & Jacobs, H. 2011. Understanding Phonology. Third Edition. Abingdon UK: Routledge.Errata in first print of 3rd edition.

Hillenbrand, J.M., Clark, M.J., Nearey, T.M. 2001. Effect of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America*, 109, 748-763.

Ingram, J. & S. Park. 1998. Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners. *Journal of the Acoustic Society of America,* 103 (2), 1161-1174.

Logan, J.S., Lively, S.E., Pisoni, D.B. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America,* 94, 1242-1255.

MacNeilage, P.F. 2011. "Sound patterns of first words and how they became linked with concepts". In: Lefebvre C (ed) *The evolution of languag*e. Cambridge: Cambridge University Press.

Magloughlin, L. (2018). */tɹ/ and /dɹ/ in North American English: Phonologization of a Coarticulatory Effect.* Doctoral Dissertation. Department of Linguistics, Faculty of Arts, University of Otawa: Canada.

Mullennix J.W, Pisoni DB, Martin CS. 1989. Some effects of talker variability on spoken-word recognition. *Journal of Acoustical Society of America*, 85, 365–378.

Odgen, R. 2009. *An Introduction to English Phonetics.* Edinburgh: Edinburgh University Press.

Peperkamp, S., Rozenn Le Calvez, J-P.N., Dupoux, E.. 2006. The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition*, 101 (3). B31–B41.

Picard, M. 2007. On Teaching the Pronunciation of Subphonemic Segments in English. *SKY Journal of Linguistics*, 20, 331–344.

Pierrehumbert, J., Talkin, D. 1992. "Lenition of /h/ and glottal stop", in G. Doherty and D. R. Ladd (eds). *Papers in Laboratory Phonology II: gesture, segment, prosody*. Cambridge, UK: Cambridge Univ. Press. 90-117.

Prator, C. H. 1971. Phonetics vs. phonemics in the ESL classroom: When is allophonic accuracy important? *TESOL Quarterly*, 5, 61–72.

Shea, C. E., Curtin, S. 2011. Experience, representations and the production of second language allophones. *Second Language Research*, 27(2), 229-250.

Sitek, K. 2011. Ipsilateral and Contralateral Phonetic Context Effects. *UC Berkeley Phonology Lab Annual Report,* 77-97.

Standwell, G. B. J. 1973. The phoneme, a red herring for language teaching? *Audio-Visual Language Journal*, 11, 119–122.

Strange, W. 1992. Learning Non-native Phoneme Contrasts: Interactions among Subject, Stimulus, and Task Variables. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.) *Speech Perception, Production Linguistic Structures*. (pp. 197-219). Tokyo: Omsha.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K. 2001. Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical society of America,* 109 (4), 1691-1704.

Trubetzkoy, N. S. 1939. *Grundzüge der Phonologie*. Travaux du cercle linguistique de Prague 7. Christiane Baltaxe, trans. *Principles of Phonology*. Berkeley: University of California Press.

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

Werker, J.F., Curtin, S. 2005. PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, 1, 197-234.

Whalen, D.H., Best, C.T., Irwin, J.R. 1997. Lexical effects in the perception and production of American English /p/ allophones, *Journal of Phonetics*, 25 (4), 501-528.

Young-Scholten, M. 1992. The Subset Principle in Interlanguage Phonology: A Preliminary Investigation. In Jonathan Leather & Allan James (eds.), pp. 200–217. Amsterdam: University of Amsterdam.

Danica Jerotijević Tišma

PERCEPCIJA I PRODUKCIJA ALOFONA KOD SRPSKIH UČENIKA ENGLESKOG KAO STRANOG JEZIKA

Rezime

Istraživanja u okviru *Modela učenja govora* (Flege 1995) kao glavnu hipotezu postavljaju da je formiranje nove kategorije u međujeziku moguće pod uslovom da učenici ciljne kontraste percipiraju kao fonetski različite od kontrasta u maternjem jeziku. Stoga je usvajanje alofona stranog jezika moguće ukoliko se razlika percipira dovoljno da se izbegne klasifikacija ekvivalencije. Isti model takođe presuponira postojanje povezanosti percepcije i produkcije bez eksplicitnog preciziranja šta čemu prethodi.

Imajući sve prethodno navedeno u vidu, rad ispituje percepciju i produkciju engleskh alofona kod srpskih učenika engleskog kao stranog jezika na tercijarnom nivou obrazovanja, tj. kako zapravo fonetski kontekst utiče na percepciju i produkciju konsonanata, jer se rad fokusira pre svega na alofone konsonanata. Empirijsko istraživanje sastojalo se od testiranja percepcije i produkcije kroz zadatak identifikacije i čitanja vezanog govora, koji je trebalo da delimično simulira situaciju u spontanom govoru. Rezultati percepcije analizirani su tako da se dobiju podaci o nivou tačnosti, dok su se u zadatku produkcije tražili podaci o nivou produkcije, odnosno koliko studenti u praksi izgovaraju alofone nalik izvornim govornicima. Ukupno 35 studenata druge godine anglistike je učestvovalo u istraživanju.

Rezultati ukazuju na diskrepantni odnos percepcije i produkcije, sa višim nivoom percepcije u većini slučajeva, ali ne bez iznenađenja, jer je za dve alofonske varijacije produkcija bila na višem nivou od percepcije. Razlike u percepciji i produkciji ukazuju na to da je proces usvajanja glasova u toku, ali takođe preispituju pretpostavljeni linerani odnos gde percepcija nužno prethodi produkciji. Rad se takođe osvrće na važne pedagoške implikacije vezane za nastavu izgovora i alofona u srpskom kontekstu učenja engleskog kao stranog jezika.

danica.tisma@filum.kg.ac.rs