

New vs. similar sound production accuracy: The uneven fight

Jolanta Sypiańska
University of Szczecin
jolanta.sypianska@usz.edu.pl

Elena-Raluca Constantin
Military Technical Academy of Bucharest

Abstract

The Speech Learning Model states that the process of equivalence classification reduces the accuracy of a similar L2 sound by forming an L1-L2 merged category whereas new sounds tend to be pronounced with greater accuracy due to lack of perceptual linkage with an L1 category (Flege 1995). We found further support for this differentiation in the production of the canonical schwa as a new sound by L1 Polish/L2 English speakers and as a similar sound by L1 Romanian/L2 English speakers, who produced an L1-L2 schwa merger. The aim of the current paper is to further investigate the production accuracy of new and similar sounds. First of all, a control group that consisted of native Romanian speakers was used to analyse the L1-L2 similar sound merger. Secondly, a measure of overall pronunciation ability based on foreign accent ratings (FAR) was included. The results confirm that production accuracy for new sounds is much greater than for similar sounds which form merged diaphones with the equivalent L1 sound. As a result, neither sound is produced on target when compared to a native speaker reference. Moreover, new sound production accuracy benefits much more from an increase in general proficiency and overall L2 pronunciation ability.

Keywords: new and similar sounds; schwa; production accuracy; foreign accent rating; level of proficiency.

1. Introduction and preliminary remarks

The Speech Learning Model (SLM) states that the process of equivalence classification reduces the accuracy of a similar L2 sound by forming an L1-L2 merged category whereas new sounds tend to be pronounced with greater accuracy due to lack of perceptual linkage with an L1 category (Flege 1995). We found further support for this differentiation in the production of the canonical

schwa as a new sound by L1 Polish/L2 English speakers and as a similar sound by L1 Romanian/L2 English speakers, who produced an L1-L2 schwa merger. The aim of the current paper is to find further support for the differences in production accuracy between new and similar sounds. In terms of similar sounds, the focus is on the L1-L2 merger that results from the process of equivalence classification. Another aim is to search for possible factors that condition new and similar sound production accuracy. The factors included in the analysis are general level of proficiency and overall pronunciation ability measured by means of foreign accent rating (FAR).

The layout of the paper is as follows: first of all, the hypotheses for category formation of similar and new sounds that derive from the assumptions of the Speech Learning Model are overviewed, hereby including the process of equivalence classification which predicts different accuracy for new and similar sounds. The outcomes which the process of equivalence classification entails for both the native and the second language sounds are subsequently explained followed by a description of how language experience impacts category assimilation and dissimilation. The assumptions of SLM are then applied to the English schwa acquired by the L1 Romanian group as a similar sound and the L1 Polish group as a new sound. Finally, the design of the study and its findings are presented.

2. New and similar sounds in SLM

This study investigates the acquisition of L2 English schwa in terms of the correspondence between the target and the L1 Romanian/L1 Polish native segments. The theoretical framework applied in the paper is Flege's (1995, 2002, 2005) Speech Learning Model (henceforth SLM). The main assumption that we derive from the framework is that the degree of phonetic similarity correlates inversely with L2 phonemes' learnability. Given the hypotheses generated by SLM, L2 segmental acquisition is envisaged as a process dependent on the interaction between the native and the target phonetic systems. A basic assumption the model builds on is that the mechanism for phonetic category formation is not constrained by age, being applied effectively in L2 learning. Moreover, sufficient exposure to the target language will finally lead to phonetic learning, i.e. changes in the perception and the production of L2 phonemes as Flege (1988: 229–230) claims. The SLM also posits that the phonetic systems responsible for both perception and production remain flexible and

upon encountering L2 sounds they can be reorganized by adding new phonetic categories or modifying existing ones (Flege 1995: 233). Phonetic evidence for the adaptability of phonetic systems is presented and thoroughly discussed in research studies concerned with nativelike oral speech production that can be achieved even by late learners (Flege & MacKay 2004).

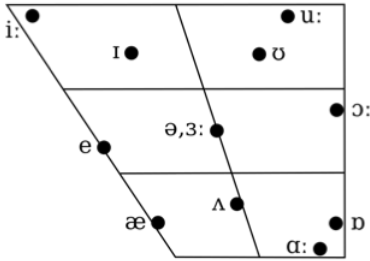
The degree of similarity between L1 and L2 phonemes is essential for the L2 sounds' learnability. The SLM maintains that the more distant an L2 sound from the closest L1 category is, the more learnable this sound is likely to be, i.e. the more likely it is that a new category for this sound will be established. The more phonetically similar an L2 sound is to an L1 category, the less likely it is that the learner will set up a relevant phonetic category. This is due to the operating cognitive mechanism of equivalence classification which hinders the formation of categories for similar sounds since they are perceived as realizations of existing L1 categories (Flege 1987). Nevertheless, this mechanism impedes, but does not fully block the formation of new categories for similar sounds. Phonetic learning is feasible without category formation because cross-language subcategorical differences are auditorily accessible to language learners (MacKay et al. 2001: 517). For speech sounds which represent one and the same phonetic category, i.e. they are subcategories, a merged category, a composite that combines the properties of the corresponding L1 and L2 speech sounds will be set up (Flege 2005).

To clarify the assimilation vs. dissimilation discrimination, the SLM proposes that when learners are unable to create a new category for an L2 vowel because it is too similar to an existing L1 vowel, the two vowels will eventually form a composite – coming to resemble each other (Flege 2005).

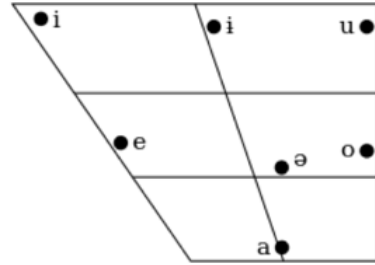
3. New and similar L2 English schwa

Considering the Romanian, Polish and English vowel plots in Figure 1, schwa stands for a case in point as the present experimental research is aimed at analysing the production of similar phones in a foreign language, with a similar L2 phone being defined (Flege 1987: 58) as an L2 phone which is realized in an acoustically different manner than an easily identifiable counterpart in L1. Since in Romanian schwa stands for a monophthong which is similar to its English counterpart, the assumption is that the L1 Romanian group will not establish a similar category to the English schwa. Polish, on the other hand,

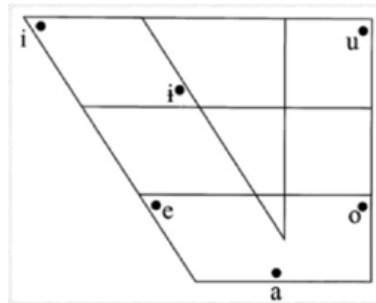
does not have a mid central vowel thus the L1 Polish group should establish a new category for the L2 English schwa.



English (Roach 2004)



Romanian (Sarlin 2014: 18)



Polish (Jassem 2003)

Figure 1. The vowel plot of English, Romanian and Polish vowels.

Sypiańska and Constantin (2018) verified how Flege's equivalence classification (1995) operates with respect to the acquisition of new and similar sounds on two groups of bilinguals (Group 1: L1 Polish, L2 English; Group 2: L1 Romanian, L2 English). The English schwa was a new sound for the L1 Polish group as Polish does not have a mid central vowel, and a similar sound for the L1 Romanian group as Romanian has a schwa vowel that is backer in comparison to the English target. The results confirmed the assumptions of the SLM as the L1 Polish group established a separate category for the L2 English schwa. The L1 Polish speakers also produced the sound more accurately with

greater proficiency. However, the L1 Romanian group showed signs of equivalence classification and produced the L2 schwa with an overall lower accuracy regardless of level of proficiency. The current paper offers an improvement of the above study design by means of adding a control group which will allow to investigate the L1-L2 merger and an analysis of how FAR conditions new and similar sound production accuracy.

3.1. The phonetics of English and Romanian schwa vowels

English schwa often stands for a weak or reduced vowel because it is the outcome of neutralization of vowel quality contrasts (Flemming 2007). Due to vowel reduction or resistance to being stressed, schwa is also commonly restricted to unstressed syllables in English. As a matter of fact, the basis for the weakness of schwa has been the subject of much research by phonologists (Van Oostendorp 2000) who made a clear-cut distinction in their recent proposals between the nature of English schwa as a mid-central vowel, and English schwa as a vowel that lacks a well-defined target, and so assimilates strongly to surrounding segments, exhibiting substantial variation in its vowel quality. Thus, recent phonetic studies conducted in the literature (Flemming 2007) indicate that both kinds of schwa vowels exist in English: a true mid central vowel and a contextually variable vowel, findings which are in line with Lass (2009).

Furthermore, Flemming and Johnson (2007) identified major phonetic differences between schwa vowels in word-final position, as in (1), and schwa vowels in other positions, as in (2).

(1) china, comma

(2) suppose, probable

(Flemming and Johnson 2007)

Word-final schwa vowels have a relatively consistent vowel quality, usually mid central, and undergo less coarticulation (Silverman 2011). Word-internal schwa vowels, on the other hand, are relatively high and vary contextually in backness and lip position. Therefore, variability of nonfinal schwa (particularly F2) was accounted for in terms of assimilation by context. Flemming (2004) argues that two main related factors are involved: word-medial schwa

is (i) very short, and (ii) it does not minimally contrast with other vowel qualities. These two factors are correlated since it is the short duration of non-final unstressed syllables that favours the neutralization of vowel quality contrasts in these contexts. In a nutshell, “to realize a particular vowel quality in a word, it is necessary to move from the articulatory position of the previous segment to the target for the vowel and then on to the position of the following segment. As the duration of the vowel decreases, it can become difficult to complete the required movements, especially if the vowel target is far from the targets for the preceding or following segments, because the articulators would have to move too fast to complete the movements in the time available” (Flemming 2007: 12).

As a matter of fact, the schwa vowel has an exceptional position, displaying a high level of context-dependency and a huge amount of variability. It is this high-context sensitivity in particular that has led to the assumption that schwa can also be targetless when not surfacing in unstressed word-final positions. Therefore, schwa is phonetically realized with an active gesture that is, however, overlapped by the gesture of the following full vowel or a vowel underspecified for tongue position (Browman et al. 1992: 26).

Basically, it is susceptibility/insusceptibility to coarticulation that makes English variable schwa in non-final word-position not be analyzed as a particular vowel quality, and the mid-central schwa found in word-final unstressed syllables be analyzed as a distinct category exhibiting a particular vowel quality.

Added to all these, it is essential for the current research to distinguish two main functions of the English schwa. Thus, as Heselwood argues (2007: 148) “two types of function are identified for schwa: anaptyctic (3), and positional (4).”

- (3) today, abbot
- (4) support, lotus
- (5) computer

Heselwood (2007: 148)

In (5), the first occurrence of schwa is anaptyctic and the second positional since a consonant sequence such as /k₁mp/ is unpronounceable and therefore needs the insertion of a support vowel (a manifestation of the phenomenon of epenthesis), whereas the final vowel segment in *computer* (this word is fol-

lowed here by a pause) functions, phonologically, as a distinctive segment (in opposition to the constitutive function of a segment) by virtue of the opposition segment position / empty position. The last schwa in (6) is also anaptyctic, since it is inserted into the sequence as a vowel sound, qualitatively assimilated to neighbouring sounds and for that reason different from theoretical canonical schwa, mainly in order to ease pronunciation.

(6) and get rid of it

(Heselwood 2007: 148)

To put it differently, anaptyctic schwa vowels are predicted to be influenced by the phonetic context to a more considerable extent than positional schwa vowels, with anaptyctic schwa vowels used as a mere type of epenthesis meant to ease the pronunciation of a consonant cluster, and positional schwas used as phonologically distinct segments.

Similarly, the Romanian schwa vowel also has an anaptyctic function beside the positional function, since targetless schwa in Romanian was also called in the literature “une voyelle vicaire/sonus vicarius” (‘a substitute vowel’). As Avram (1990: 9) claims, “en tenant compte de la ‘fonction’ qui vient d’être mentionnée et aussi d’autres particularités des voyelles [ə] et [i], Sextil Puşcariu a nommé ces deux sons des *voyelles vicaires*” (‘bearing in mind the function and other peculiarities of the [ə] and [i], they may be called substitute vowels’).

Let us consider the examples in (7a–c) below.

(7a) *alt[ə]cine* ‘another’

(7b) *opt[i]sprezece* ‘eighteen’

(7c) *ours[ə]blanc* ‘a polar bear’

(Avram 1990: 9)

The Romanian central vowels in (7a), (7b) respectively, play the same role as their French counterpart, namely ease of articulation (Avram 1990: 9). According to Avram (1990: 9), the insertion of a parasite vowel in consonant clusters that are difficult to utter is a frequent phonological phenomenon irrespective of the occurrence of the parasite vowel with respect to the morphemic boundaries.

To avoid too much variation brought about by contextual coarticulation, our experiment is aimed at investigating the latter type of schwa, that is, positional schwa, a distinct segment on its own with inherent vowel consistency as proven insofar.

Unlike the English schwa, the Romanian schwa is consistent with both stressed and unstressed positions, with some differences in F2 between the two occurrences. Table 1 illustrates the mean formant values of the Romanian schwa (phonetically transcribed as /ʌ/) and other Romanian vowels, for each gender, in each stress condition (Renwick 2012: 158).

Table 1. Mean standard deviations (Hz) for the first and second formants for vowel tokens pooled across speakers, separated by gender and stress condition (14 female speakers; 3 male speakers; non-normalized data) (Renwick¹ 2012: 158).

			/a/	/e/	/i/	/ɨ/	/o/	/u/	/ʌ/
F	F1	STR	897	603	377	444	591	411	636
		UNS	856	552	333	450	573	406	583
	F2	STR	1463	2095	2720	1600	1003	1106	1503
		UNS	1473	1961	2745	1850	1073	1324	1595
M	F1	STR	679	495	317	381	497	363	519
		UNS	685	438	294	392	494	365	490
	F2	STR	1302	1737	2151	1482	993	1116	1377
		UNS	1239	1710	2149	1683	986	1232	1446

One of the findings reported by Renwick (2014: 101) is that the Romanian schwa vowel occupies its own acoustic space, and does not exhibit the degree

¹ Renwick (2014: 13) takes a position on the transcription conventions for each monophthong in the Romanian L1 vowel inventory, with the specification that the Romanian mid central schwa is transcribed /ʌ/ in their studies, in line with Steriade (2008), unlike the standard transcription of the schwa vowel /ə/. The latter transcription is substituted with the one encapsulated in Table 1 since the usage of the standard phonetic transcription may suggest that the schwa vowel in question is a reduced one, which perhaps participates in phonological neutralizations in prosodically weak positions. As Renwick (2014: 13) reports, while the mid central vowel historically developed in unstressed syllables and was likely a reduced vowel, it functions synchronically as a full vowel and should be distinguished from reduced vowels.

of variability of formant values that is expected from a targetless reduced vowel. The data analyzed in Renwick (2014: 101) obviously show that durationally, the Romanian L1 schwa vowel patterns with other vowels of its height class and thus is not reduced in dimension. For these reasons, the Romanian L1 schwa vowel should be treated unambiguously as a full vowel, even in unstressed positions where it is contrastive as a morpheme (Anghelina 2008: 529).

Moreover, as Chițoran (2002: 210) points out the alternation of stressed [a] with unstressed schwa in Romanian may be interpreted as an instance of vowel reduction, similar to that typical of English, for instance. Nevertheless, schwa does not stand for a phonologically reduced vowel in all contexts in Romanian, and it has phonemic status. It may itself surface under stress, and there is evidence of underlying schwa in various forms. The roots in (8a) for example, all feature underlying schwa, which surfaces in both stressed and unstressed positions. The schwa vowels in (8b) do not alternate with any other vowels, and their occurrence is unpredictable, thus supporting the view that they are part of the underlying representation of these forms. Also, the minimal pairs in (8c) bring further evidence in favour of the Romanian schwa vowel as an underlying phoneme.

(8a) *mătură* /'məturə/ 'broom' – *măturică* /mətu'rikə/ 'broomie' (dim.)
pătură /'pəturə/ 'blanket' – *păturică* /pətur'ikə/ 'blankie' (diminutive)

(8b) *pământ* /pə'mint/ 'earth'
părinte /pə'rinte/ 'parent'
grătar /grə'tar/ 'grill'

(8c) *păr* /pər/ 'hair'
par /pər/ 'pole'

(Chițoran 2002: 210)

Reduction alternations are common in Romanian (9), since Romanian exhibits phenomena of derived environment effects of stressless vowel reduction (Khanjian 2009: 185). Thus, in Romanian, [á] turns into [ʌ] when stress shifts (Steriade 2008).

(9a) *sărac* /sʌ'rak/ 'poor' – *sărăcuț* /sʌrʌk-'uts/ (diminutive) 'poor'

(9b) *papuc* /pa'puk/ 'slipper' – *papucel* /papuʃ-'el/ (diminutive) 'slipper'

Furthermore, given the Romanian vowel space in Figure 1 (Sarlin 2014: 18), we may notice that the Romanian schwa is backer when compared to the English positional schwa and this is a difference to take into consideration while analysing the tokens in the current paper.

In brief, the schwa categorization carried out by Oostendorp (1998) and subsequently adopted by Veloso (2007: 55) constitutes an inspiration for the current research. As shown in examples (1–9) provided in this section, such categorization includes three main types of schwa to be found both in English and Romanian: ‘e-schwa’ or epenthetic schwa – schwa that results from epenthesis which usually alternates with zero; ‘r-schwa’ or vowel-reduction schwa – schwa that alternates with a full vowel as the result of vowel reduction, and ‘s-schwa’ or stable schwa which is schwa present at the underlying representation. This type of schwa is neither the result of epenthesis nor vowel reduction. Given its lack of variation, it is stable schwa to be investigated with experimental data in this paper.

4. Methodology and experimental conditions

The current study is a continuation of previous research. Thus, the methodology is partly based on Sypiańska and Constantin (2018).

4.1. Aim

This study is a continuation of the research on the production accuracy of new and similar sounds. In particular, it offers a closer look at the behaviour of similar sound categories in bilinguals with an investigation of the L1-L2 category merger of the Romanian schwa in the L1 Romanian, L2 English bilingual group. Since level of proficiency measured with an overall placement test was found not to influence production accuracy, another aim of the current study is to verify whether overall pronunciation ability based on foreign accent ratings (FAR) conditions the accuracy with which similar sounds are produced.

4.2. Bilingual participants

An aggregate of 33 Romanian (5 females and 28 males; mean age = 21) and 25 Polish respondents (18 females and 7 males, mean age = 21) were selected for the oral production experiment, with respectively Romanian or Polish as

the L1 and English as the L2. The speakers also reported knowing other foreign languages, typically German, however with no current exposure to the language and negligible proficiency.

All subjects were asked to fill in a thorough background questionnaire which elicited the participants' phonetic and linguistic background. More precisely, the informants were asked to provide personal data with respect to the number of foreign languages they spoke, the onset age and the end age for each target language. They were also required to report the number of hours in the target language they benefited from both in and outside the classroom. Other main issues such as access to institutionalized phonetic training and any long-term stays in the target language country were also included in the questionnaire. It is to note that no participant within the two groups reported language or hearing impairment.

It was compulsory for all the informants to sit for an ERASMUS+ online language test since they all joined mobilities abroad in various receiving countries. The language test complied with the European Language Passport, which is a standardized template for self-assessment of language skills aimed at testing non-natives' reading comprehension abilities, listening comprehension abilities, grammar and vocabulary skills as well as performance with respect to key communicative phrases. On the basis of the test results, three levels of proficiency were considered in English as an L2. Thus, 11 L1 Romanian respondents had a B1 level of English, 13 L1 Romanian speakers were assigned a B2 level of English and the other 9 subjects a C1 command of English. The L1 Polish respondents were assigned to B1 (N = 9), B2 (N = 8) and C1 (N = 8).

4.3. Oral production stimuli

All informants were required to read words in citation form aloud. The stimuli were two-syllable words with word-final schwa preceded by a voiceless obstruent. This was the phonetic environment adopted so that any contextual variation or coarticulation would be minimised. The Romanian group was asked to read Romanian and English words once. The Polish group read only English words once. Twenty-four fillers were inserted for distraction. A word list of the recorded schwa vowel tokens may be found in Table 2. The target vowel encapsulated in Table 2 is in bold.

Table 2. Word list featuring word-final schwa vowel tokens in English and Romanian.

	Transcription	Orthography	Gloss	Stress
1.	/ˈsi:kə/	seeker	–	UNS
2.	/ˈfatə/	fată	girl	UNS
3.	/ˈti:tʃə/	teacher	–	UNS
4.	/ˈkapə/	capă	cloak	UNS
5.	/kəmˈpjʊ:tə/	computer	–	UNS
6.	/ˈfrikə/	frică	fear	UNS
7.	/ˈbitə/	bitter	–	UNS
8.	/ˈpatə/	pată	stain	UNS
9.	/ˈflæʃə/	flasher	–	UNS
10.	/ˈsəpə/	sapă	shovel	UNS
11.	/ˈkɪsə/	kisser	–	UNS
12.	/ˈmɑrfə/	marfă	goods	UNS

(Sypiańska and Constantin: 2018)

4.4. Procedure

The recordings were carried out in a sound-treated booth at Ferdinand I Military Technical Academy of Bucharest and Adam Mickiewicz University in Poznań (Poznań: Røde NT1-KIT condenser microphone connected to a computer via a Focusrite Scarlett 2i2 2Gen audio interface; Bucharest: Vivanco DM 97 microphone connected to a laptop computer via a t.c.electronic connect 8). All participants were explicitly instructed to avoid rhotacisation of the final vowel, which could have altered the values of the formants needed for the present study. All rhotacised tokens that were subsequently identified in the speech samples were removed. Then F1 and F2 values were extracted from the vowel midpoint by means of a Praat script (Lennes 2003) and were normalized by means of the Lobanov transform (1971).

4.4.1. Procedure: L1-L2 merger

In order to further investigate the L1-L2 merger of the schwa category produced by the bilingual L1 Romanian, L2 English group, a control group was

introduced into the study design (Table 3). The control group consisted of 28 native speakers of Romanian who reported no exposure to foreign languages in the 10 years prior to the study. There were 25 females and 3 males with a mean age of 50. The mean age of the control group was much higher than the research group, however, it is not possible to find an aged matched group with no exposure to foreign languages in the Romanian society. Since no reports of changes in the Romanian schwa sound have been reported in the literature on the Romanian language, it was assumed that no exposure to foreign language is more important than age matching for this particular study.

Two one-way ANOVAs were run to test the effect of the independent variable of Group (Control; Bilingual) on the dependent variables of F1 and F2 of the Romanian schwa vowel.

Table 3. Summary of participant information for the L1-L2 merger procedure.

Groups	N	Mean age	Sex	Tested element
Control	28	50	M=3; F=25	Romanian schwa (N=336)
Research	33	21	M=28; F=5	Romanian schwa (N=396)
Total	47			

4.4.2. Procedure: Overall pronunciation ability

Our initial findings showed a beneficial influence of general level of proficiency on the accuracy with which new sounds are produced. However, no such impact on similar sound production accuracy was noted. Bearing this in mind, the question for the current study was whether there is a difference in how overall pronunciation ability conditions similar and new sound production accuracy.

The recordings of the word-final English schwa tokens from the previous study were given to judges for accent rating. The ratings were carried out on tokens pronounced by two groups of bilingual speakers (Group 1: L1 Polish, L2 English; Group 2: L1 Romanian, L2 English). The two groups were divided into three levels of proficiency B1, B2, C1 in their L2 English (Group 1: B1 n = 9, B2 = 8, C1 = 8; Group 2: B1 = 11, B2 = 13, C1 = 9). All the speakers were taught British English with the use of textbooks and other materials based on the British model. All the judges were phonetically trained phoneticians who

have experience in teaching and assessing the pronunciation of learners of English in an academic environment. The L1 Polish group was assessed by Polish native speakers ($N = 4$) and the L1 Romanian group was judged by Romanian native speakers ($N = 4$). This step was agreed upon as a means of minimising the bias of having Polish native speakers as judges rating the accent of the L1 Romanian group or Romanian native speakers rating the Polish accent in English as they may be more prone to harsher judges due to lack of familiarisation with the accent. English native speakers were not engaged as raters because of a generally greater familiarisation with the Polish accent rather than Romanian accent in English among inhabitants of Great Britain. The stimulus for rating was provided in the form of separate words following Flege and Munro (1994) and Moyer (1999) in a randomised data set in order to avoid harsher judgments with increased familiarisation with the set (Flege & Fletcher 1992; Munro & Derwing 1994). During the rating task, the judges were wearing headphones and listened to the tokens as many times as needed to decide on the rating. The judges were instructed to use a 9-point Likert scale (e.g. Munro & Derwing 1999; Munro, Derwing & Flege 1999; Munro & Derwing 2001) in which 1 stood for native-like and 9 meant heavy foreign accent.

In order to test the hypotheses, a General Linear Model was run with Group, FAR and Level of proficiency as main effects. Since different effects of Level of proficiency and FAR on group were foreseen, the following interaction effects were taken into consideration: Group*FAR and Group*Level of proficiency. The dependent variables included F1 and F2 of the L2 English wordfinal schwa. The Bonferroni correction for multiple comparisons was applied in order to avoid false positive results.

Table 6. Summary of participant information for overall pronunciation ability procedure.

Groups	LoP	N	Mean age	Sex	Tested element
L1 Polish, L2 English	B1 = 9; B2 = 8; C1 = 8	25	21	M=18; F=7	English schwa
L1 Romanian, L2 English	B1 = 11; B2 = 13; C1 = 9	33	21	M=28; F=5	English schwa
Total		58			

4.5. Results

In the results section, the presentation of L1-L2 merger data is followed by the results of the overall pronunciation ability.

4.5.1. Results: L1-L2 merger

The research group produced the Romanian schwa vowel with a mean F1 of 484 Hz and a mean F2 of 1550 Hz compared to 502 Hz and 1424 Hz for the controls (Fig. 2). The results of the one-way ANOVA showed no statistically significant difference for F1 between the two groups ($F = 0.00461$, $p = 0.9459$) (Table 4), however, the difference with regard to F2 reached statistical significance ($F = 6.80601$, $p = 0.00935$) (Table 5).

Table 4. One-way ANOVA results for F1 between Control and Research groups.

ANOVA F1

Source of Variation	d.f.	SS	MS	F	p-value
Between Groups	1	54.88199	54.88199	0.00461	0.9459
Within Groups	514	6 121 315.27893	11 909.17369		
Total	515	6 121 370.16092			

Table 5. One-way ANOVA results for F2 between Control and Research groups.

ANOVA F2

Source of Variation	d.f.	SS	MS	F	p-value
Between Groups	1	235 792.26125	235 792.26125	6.80601	0.00935
Within Groups	514	17 807 374.98842	34 644.69842		
Total	515	18 043 167.24967			

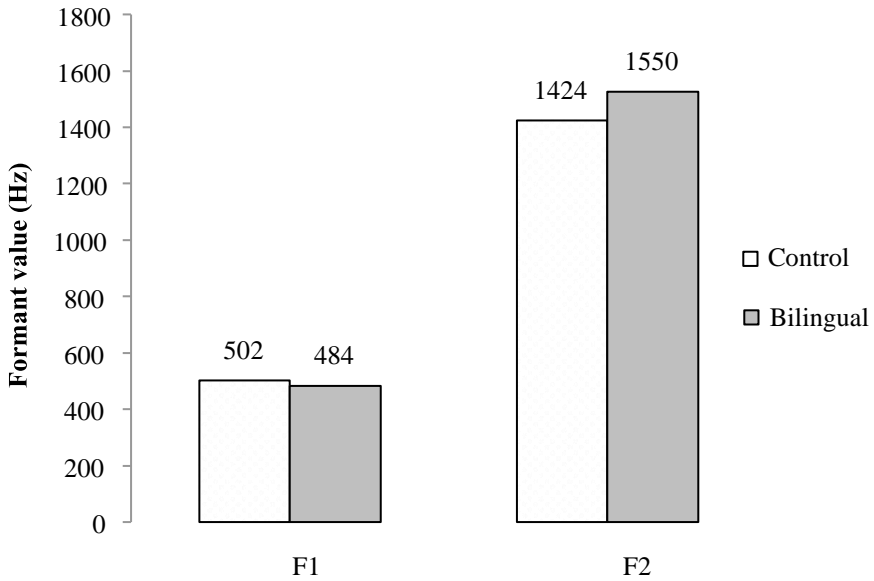


Figure 2. Mean F1 and F2 results for the Research and Control groups.

All together, the Romanian research group produced a fronter Romanian schwa than the controls. A comparison with the native English speaker reference based on data from Flemming and Johnson (2007) indicates that the bilinguals are influenced by English in their native schwa production and make the Romanian schwa fronter under the influence of the fronter English schwa (Fig.3).

4.5.2. Results: Overall pronunciation ability

As our initial findings indicated, there were differences between the L1 Polish group and the L1 Romanian group (Table 7). A full report of the General Linear Model is available in Appendix 1.

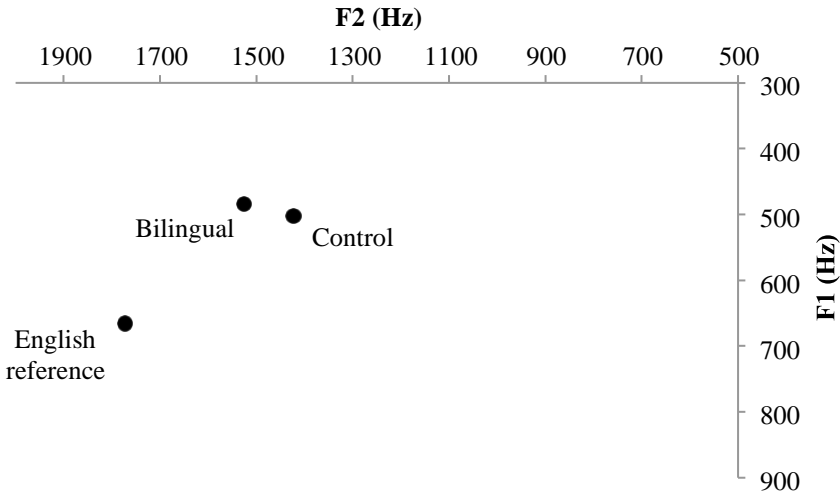


Fig. 3. Romanian schwa as produced by the Control and Bilingual group compared against the English schwa reference values based on Flemming and Johnson (2007).

Table 7. Descriptive statistics of F1 and F2 of the L2 English schwa as produced by Group 1: L1 Polish and Group 2: L1 Romanian (Sypiańska and Constantin 2018).

Groups	Sample size	Mean F1	SD F1	Mean F2	SD F2	Mean FAR
L1 Polish L2 English	121	670	14.43	1618	22.11	5.05
L1 Romanian L2 English	198	484	8.81	1550	61.54	6.87

The results showed a main effect of LoP for both F1 ($F = 41.363$, $p < .000$) and F2 ($F = 53.399$, $p < .000$). The estimated marginal means for both formants increased with every level of proficiency making the schwa vowel more English-like in height and backness (Fig. 4 and 5).

When it comes to the main effect of FAR, it reached statistical significance for the second formant ($F = 8.176$; $p < .001$) but not for F1 ($F = 0.1$; $p = 0.752$). Lower FAR scores were more likely awarded to those participants who produced an English schwa with a higher F2 that is more English-like (Fig. 6). Judging by the effect size, this tendency was rather small ($\eta^2 = 0.029$).

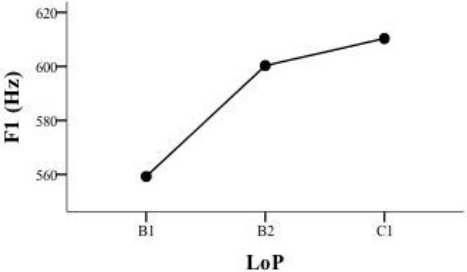


Fig. 4. Estimated marginal means for F1 according to three levels of proficiency B1, B2 and C1.

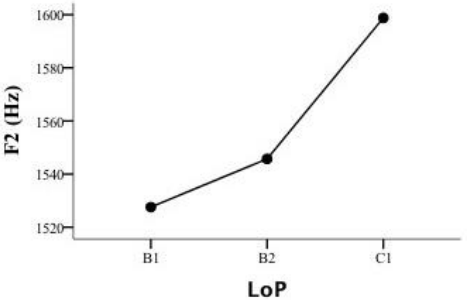


Fig. 5. Estimated marginal means for F2 according to three levels of proficiency B1, B2 and C1.

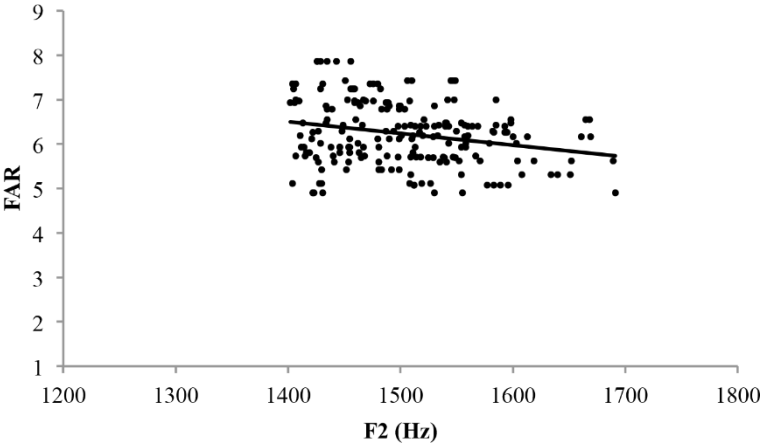


Fig. 6. Interaction effect of Group and FAR.

The most interesting results are derived from the interaction effects. There was a large interaction effect of Group*LoP for both F1 ($\eta^2 = 0.183$; $F = 41.363$, $p < .000$) and F2 ($\eta^2 = 0.234$; $F = 53.399$, $p < .000$). Both formants were higher with every LoP in the L1 Polish group but not in the L1 Romanian group (Fig. 7 and 8).

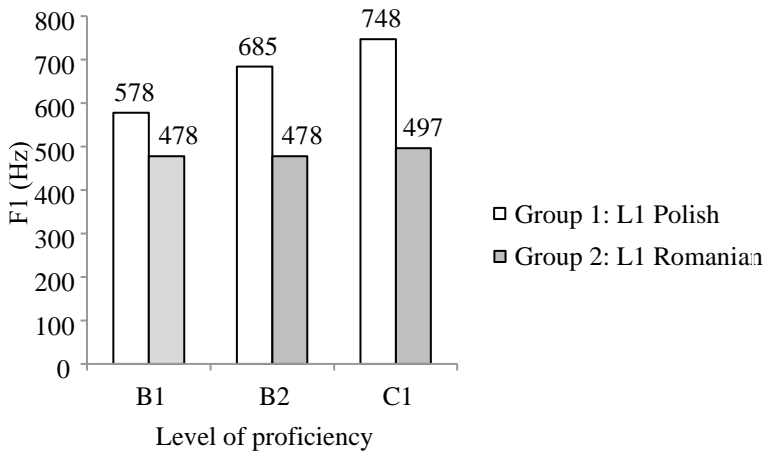


Fig. 7. Interaction effect of Group and LoP for F1.

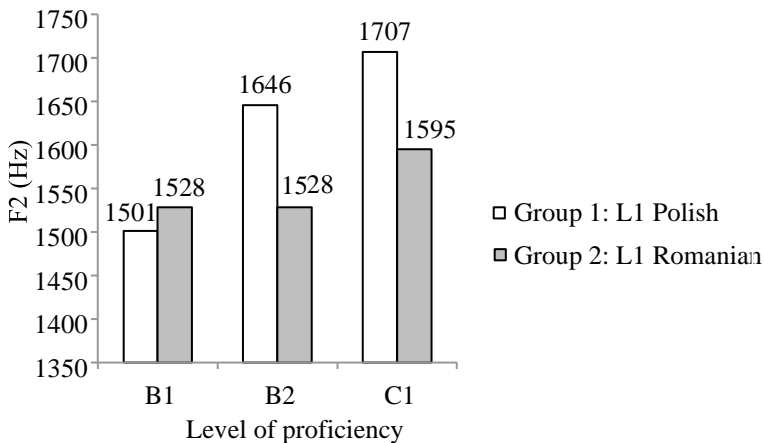


Fig. 8. Interaction effect of Group and LoP for F2.

The interaction effect of Group*FAR was much smaller but reached statistical significance for F2 ($\eta^2 = 0.039$; $F = 11.416$, $p < .001$). The ratings of overall pronunciation ability were lower with an increase of F2 in both groups but the tendency was greater in the L1 Polish, L2 English group where the trend line is much steeper (Fig. 9).

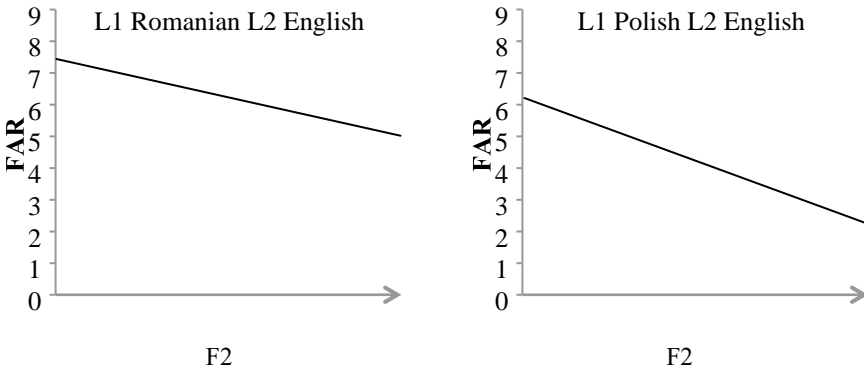


Fig. 9. Interaction effect of Group and FAR for F2.

4.6. Discussion

Our previous findings indicated that production accuracy is different for new and similar sounds. The new L2 English schwa was produced more accurately by L1 Polish speakers than the similar English schwa produced by L1 Romanian speakers. Also, the L1 Polish speakers benefited more from a higher general level of proficiency as more proficient speakers produced the L2 English schwa much more accurately than did more proficient L1 Romanian speakers. The current paper was aimed at further investigating similar sound accuracy. By adding a control group of Romanian native speakers with no exposure to foreign languages in the 10 years prior to the study, we investigated the merger that takes place between the similar sound in the L2 and an equivalent L1 sound. Although our previous research showed that the Romanian group basically pronounced both their English and Romanian schwas with almost exactly the same values, these values are also significantly different from the Romanian control. In accordance with the predictions of the Speech Learning Model

(Flege 1995), the L1-L2 merger that takes place between the similar L1 and L2 sounds results in a diaphone whose values are neither those of the L1 nor of the L2 native speaker references. In the current study, the L1-L2 diaphone is different from both the English schwa with reference values from Flemming and Johnson (2007) and the Romanian schwa produced by the controls (Fig. 10). By counting the Euclidean distances, we can further say that the diaphone is closer to the latter (Table 8).

Table 8. Euclidean distances between the diaphone and the Romanian and English schwa values.

Sound pairs	Euclidean distance
Diaphone vs. Romanian schwa	170
Diaphone vs. English schwa	312

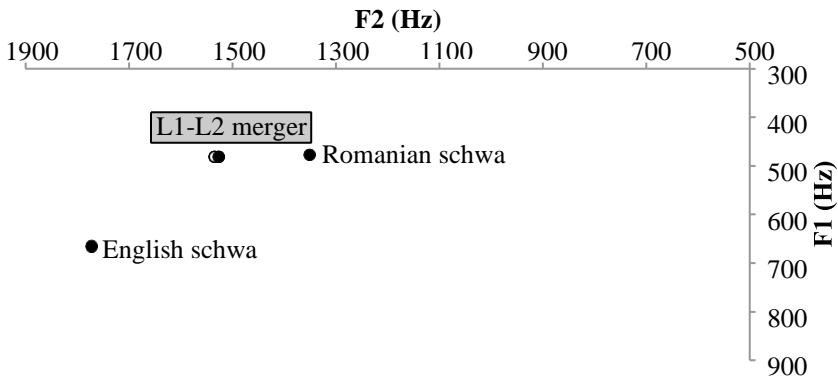


Fig. 10. L1-L2 merger.

In an attempt to find a factor which would condition both similar and new sound production accuracy, we introduced Foreign Accent Ratings that constitute a measure of general pronunciation ability. The results of the current paper show that FAR unlike general level of proficiency is connected to both new and similar sound production accuracy, but it conditions the former to a greater extent than the latter.

To sum up, bilinguals set up a phonetic category for a similar sound by producing an L1-L2 merger. The diaphone that is the result of the merger is much closer to the L1 than to the L2 reference values for the particular vowels that constitute the merged diaphone. Not only are new sounds more accurately produced, the process of new sounds acquisition benefits much more from an increase in both general L2 proficiency and overall L2 pronunciation ability measured by foreign accent ratings.

4.7. Conclusions

All in all, accuracy in production depends on the type of L2 sound. New sounds are produced more accurately. General level of proficiency conditions new and similar sound production accuracy differently as only new sounds benefit from progress in the L2. An important finding is that similar sound categories may become more accurate with the improvement of general pronunciation abilities in the L2 but the tendency is much greater for new sounds. Moreover, phonetic categories set up for similar sounds can take the form of L1-L2 mergers which are closer to native control values than L2 target categories.

Appendix 1

GLM results.

Test of Between-Subjects Effects	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	η^2
Corrected Model	F1	1482425.481 ^a	7	211775.069	40.301	0.000	0.504
	F2	1610394.067 ^b	7	230056.295	33.354	0.000	0.456
Intercept	F1	492098.329	1	492098.329	93.648	0.000	0.252
	F2	2048054.234	1	2048054.234	296.931	0.000	0.516
Group * LoP	F1	326534.151	2	163267.076	31.07	0.000	0.183
	F2	584194.681	2	292097.34	42.349	0.000	0.234
Group * FAR	F1	15716.193	1	15716.193	2.991	0.085	0.011
	F2	78737.91	1	78737.91	11.416	0.001	0.039
LoP	F1	434702.21	2	217351.105	41.363	0.000	0.229
	F2	736625.611	2	368312.805	53.399	0.000	0.278
Group	F1	32.87	1	32.87	0.006	0.937	0
	F2	19271.326	1	19271.326	2.794	0.096	0.01
FAR	F1	527.033	1	527.033	0.1	0.752	0
	F2	56391.993	1	56391.993	8.176	0.005	0.029
Error	F1	1460830.169	278	5254.785			
	F2	1917481.377	278	6897.415			
Total	F1	99474082	286				
	F2	678738191	286				
Corrected Total	F1	2943255.65	285				
	F2	3527875.444	285				

^a R Squared = .504 (Adjusted R Squared = .491)

^b R Squared = .456 (Adjusted R Squared = .443)

References

- Avram, Andrei. 1990. Remarques sur les voyelles neutres du roumain et du français. *Word* 24(1–3). 8–13.
- Anghelina, Cătălin. 2008. On some of the uses of schwa (ă) in Romanian verbal morphology. *RRL* LIII. 529–532. București.

- Browman, Carol & Louis Goldstein. (1992). "Targetless" schwa: An articulatory analysis. In John Kingston & Mary E. Beckman (eds.), *Papers in laboratory phonology ii*, 26–56. Cambridge: Cambridge University Press.
- Chițoran, Ioana. 2002. *The phonology of Romanian: A constraint-based approach*. New York: Mouton de Gruyter.
- Flege, James Emil. 1987. The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15. 47–65.
- Flege, James Emil. 1988. The production and perception of foreign languages. In Harris Winitz (ed.), *Human communication and its disorders*, 224–401. Norwood, NJ: Ablex.
- Flege, James Emil. 1995. Second language speech learning: Theory, findings, and problems. In Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 233–276. Timonium, MD: York Press.
- Flege, James Emil. 2002. Interactions between the native and second-language phonetic systems. In Petra Burmeister, Thorsten Piske & Andreas Rohde (eds.), *An Integrated View of Language Development: Papers in Honor of Henning Wode*, 217–244. Trier: Wissenschaftlicher Verlag.
- Flege, James Emil. 2005. Origins and development of the Speech Learning Model. Univ. of Alabama at Birmingham Keynote lecture presented at the 1st ASA Workshop on L2 Speech Learning Simon Fraser Univ., Vancouver, BC April 14–15, 2005.
- Flege, James Emil & MacKay, Ian. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition* 26, 1–34.
- Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Bruce Hayes, Robert Kirchner & Donka Steriade (eds.), *Phonetically-based Phonology*, 232–276. Cambridge: Cambridge University Press.
- Flemming, Edward. 2007. *The Phonetics of schwa vowels*. Retrieved 20th May at <http://web.mit.edu/flemming/www/paper/schwaphonetics.pdf>.
- Flemming, Edward & Sarah Johnson. 2007. Rosa's roses: reduced vowels in American English. *Journal of the International Phonetic Association* 37, 83–96.
- Heselwood, Barry. 2007. Schwa and the phonotactics of RP English. *Transactions of the Philological Society* 105(2). 148–187.
- Jassem, Wiktor. 2003. Illustrations of the IPA: Polish. *Journal of the International Phonetic Association* 33(1). 103–107.
- Khanjian, Hrayr. 2009. Stress Dependent Vowel Reduction. The Berkeley Linguistics Society 35th Annual Meeting, February 14–16 2009. Berkeley California. No 1.
- Lass, Robert. 2009. On schwa: synchronic prelude and historical fugue. Donka Minkova (ed.), *Phonological weakness in English: From Old to Present-Day English*, 47–77. Basingstoke and New York: Palgrave Macmillan.
- Lennes, Mette. 2003. Praat script: Get pitch formants and duration from labeled segments.
- Lobanov, Boris M. (1971). Classification of Russian vowels spoken by different speakers, *Journal of the Acoustical Society of America*. 49, 606–608.

- MacKay, Ian, Flege, James Emil, Piske, Thorsten & Carlo Schirru. (2001). Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America* 110(1). 516–528.
- Oostendorp, Marc van. (2000). *Phonological Projection; A Theory of Schwa in Optimality Theory*. Berlin: Mouton De Gruyter.
- Renwick, Margaret Elspeth Lambert. 2012. *Vowels of Romanian: historical, phonological and phonetic studies*. Ithaca, NY: The Cornell University PhD dissertation.
- Renwick, Margaret Elspeth Lambert. 2014. *The Phonetics and Phonology of Contrast*. Berlin: Mouton De Gruyter.
- Roach, Peter. 2004. British English: Received Pronunciation. *Journal of the International Phonetic Association* 34(2). 239–245.
- Sarlin, Mika. 2014. *Romanian grammar*, 2nd ed. Helsinki: Books on Demand GmbH.
- Silverman, Daniel. 2011. Schwa. Marc van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.), *The Blackwell Companion to Phonology*, 628–642. Blackwell Publishing.
- Steriade, Donca. 2008. *Contour correspondence: The segmental evidence*. 11th International Symposium on Chinese Languages and Linguistics, National Chiao Tung University, Taiwan.
- Sypiańska, Jolanta & Constantin, Elena-Raluca. 2018. *Phonetic learnability and accuracy of new and similar L2 sounds*. 12th International Conference on Native and Non-native Accents of English. University of Łódź.
- Veloso, João. 2007. Schwa in European Portuguese: The Phonological Status of [i]. In Olivier Crouzet & Jean Pierre Angoujard (eds.), *Actes des/Proceedings of JEL'2007. Schwa(s). 5.èmes Journées d'Etudes Linguistiques*, 55–60. Nantes: Université de Nantes.

Corresponding author:

Jolanta Sypiańska
University of Szczecin
Instytut Językoznawstwa
al. Piastów 40B
71-065 Szczecin
Poland
jolanta.sypianska@usz.edu.pl