


Classification of English vowels in terms of Cypriot Greek categories: The role of acoustic similarity between L1 and L2 sounds

GEORGIOS P. GEORGIOU 

Department of Languages and Literature, University of Nicosia, Nicosia, Cyprus

Director of the Phonetic Lab, University of Nicosia, Nicosia, Cyprus

georgiou.georg@unic.ac.cy

Abstract

Previous evidence has suggested that acoustic similarity between first language (L1) and second language (L2) sounds is an accurate indicator of the speakers' L2 classification patterns. This study investigates this assumption by examining how speakers of an under-researched language, namely Cypriot Greek, classify L2 English vowels in terms of their L1 categories. The experimental protocol relied on a perception and a production study. For the purpose of the production study, two linear discriminant analysis (LDA) models, one with both formants and duration (FD) and one with only formants (F) as input, were used to predict this classification; the models included data from both English and Cypriot Greek speakers. The perception study consisted of a classification task performed by adult Cypriot Greek advanced speakers of English who permanently resided in Cyprus. The results demonstrated that acoustic similarity was a relatively good predictor of speakers' classification patterns as the majority of L2 vowels classified with the highest proportion were predicted with success by the LDA models. In addition, the F model was better than the FD model in predicting the full range of responses. This shows that duration features were less important than formant features for the prediction of L2 vowel classification.

Keywords: classification, acoustic similarity, Cypriot Greek, English, vowels

Résumé

Des études antérieures ont suggéré que la similarité acoustique entre les sons de la première langue (L1) et ceux de la deuxième langue (L2) est un indicateur précis des schémas de

The study was conducted with the support of the Phonetic Lab and the Cyprus Linguistics and Humanities Research group of the University of Nicosia. I would like to thank all participants and the anonymous reviewers.

classification des locuteurs dans la L2. Cet article étudie cette hypothèse en examinant comment les locuteurs d'une langue peu étudiée, à savoir le grec chypriote, classent les voyelles anglaises de la L2 en fonction de leurs catégories de la L1. Le protocole expérimental repose sur une étude de perception et une étude de production. Pour l'étude de la production, deux modèles d'analyse discriminante linéaire (LDA), l'un avec les formants et la durée (FD) et l'autre avec seulement les formants (F) comme entrée, ont été utilisés pour prédire cette classification; les modèles comprenaient des données provenant de locuteurs anglais et grecs chypriotes. L'étude de perception a consisté en une tâche de classification réalisée par des adultes chypriotes grecs avec un niveau avancé d'anglais et résidant de manière permanente à Chypre. Les résultats ont démontré que la similarité acoustique était un prédicteur relativement bon des modèles de classification des locuteurs, car la majorité des voyelles L2 classées avec la proportion la plus élevée ont été prédites avec succès par les modèles LDA. En outre, le modèle F fonctionnait mieux que le modèle FD pour prédire l'ensemble des réponses. Cela démontre que les caractéristiques de durée sont moins importantes que les caractéristiques de formants pour la prédiction de la classification des voyelles L2.

Mots clés: classification, similarité acoustique, grec chypriote, anglais, voyelles

1. INTRODUCTION

The perception of second language (L2) sounds is based to a great extent on the speakers' first language (L1) (Darcy and Krüger 2012; Grimaldi et al. 2014; Kartushina and Frauenfelder 2014; Georgiou 2021b, 2022a; Melnik-Leroy et al. 2022). The dominant psycholinguistic approaches argue that L2 sounds are mapped as exemplars of existing L1 categories (Best 1995; Flege 1995, 2003; Best and Tyler 2007). This often impedes speech perception in the L2, especially when two or more L2 sounds are mapped to a single L1 category. A possible explanation is that as a result of the alteration of lower-level perceptual processing due to L1 experience, higher-level processing, which includes the development of L2 representations, is also altered (Iverson et al. 2003). For instance, Baigorri et al. (2019) found that Spanish-English late bilinguals had poor discrimination accuracy over American English / Λ – æ / and / Λ – a / vowel contrasts since both English vowels were assimilated to Spanish vowel category / a /. Furthermore, Georgiou et al. (2020a) showed that Russian speakers with either a large or small vocabulary in English struggled to discriminate English / e – æ / vowel contrast as both vowel members were perceived as acoustically similar instances of Russian / e / vowel.

The *size* and *complexity* of L1 and L2 phonological systems are believed to be crucial for the perception of L2 sounds (Fox et al. 1995; Hacquard et al. 2007; Iverson and Evans 2007, 2009; Souza et al. 2017). For example, Georgiou (2021a) argued that Cypriot Greek learners of Italian discriminated the Italian [e – ɛ] contrast only to a moderate extent. This can be explained by the fact that Italian has a larger vowel system compared to Cypriot Greek, containing vowel qualities that do not create contrast in the speakers' L1. In addition, Iverson and Evans (2007) found that speakers of Norwegian and German, two languages with extended and complex vowel systems, were more successful in identifying English vowels than

speakers of French and Spanish, two languages with smaller vowel systems. A further important finding of the study was that speakers of all four languages relied on the first two formant frequencies, formant movement, and duration during speech perception, although the latter two features are absent from Spanish and French. In contrast to the previous study, evidence from Spanish and Australian English listeners of Brazilian Portuguese suggested that the two groups of listeners did not differ in the discrimination of non-native vowel contrasts, although Spanish has a smaller and simpler vowel system than Brazilian Portuguese whereas Australian English's system is larger and more complex (Elvin et al. 2014). All these lead to the assumption that vowel size and complexity might not always be good predictors of non-native speech perception, as acoustic-phonetic features also play a role. It has been suggested that *acoustic similarity* of L1 versus L2 sounds may predict the L2 speech perceptual patterns.

There is either direct or indirect recognition from several theoretical speech models (e.g., Perceptual Assimilation Model (Best 1995), Speech Learning Model (Flege 1995), Perceptual Assimilation Model-L2 (Best and Tyler 2007), Second Language Linguistic Perception Model (Escudero 2009), Speech Learning Model-r (Flege and Bohn 2021), Universal Perceptual Model (Georgiou 2021a)) that the phonological and articulatory-phonetic or acoustic-phonetic similarity between native and non-native sounds is a successful predictor of the listeners' non-native sound perceptual patterns. Acoustic similarity can be measured with a variety of techniques including *Linear Discriminant Analysis* (LDA) (Klecka 1980), which aims to assess how well the non-native vowels fit with the center of gravity of the input corpus tokens, giving a predicted probability on how each non-native vowel will be mapped into the learners' L1 categories (Elvin et al. 2021). The L1 models are trained on measurements collected from natural speech such as the first (*F1*), second (*F2*), and third formants (*F3*), durations, etc., and subsequently the corresponding non-native test measurements are supplied to the trained models to form the predictions. An LDA model was used by Strange et al. (2005) to examine the perceptual similarity of North German front rounded vowels and American English front unrounded and back rounded vowels. The authors concluded that perceptual similarity between the vowels of the two languages could not always be predicted successfully by acoustic similarity. However, the results of Gilichinskaya and Strange (2010) provide evidence of the success of this technique, implying that crosslinguistic acoustic similarity can predict non-native sound categorization. They found that the assimilation of American English vowels by inexperienced Russian listeners could be predicted for all but one vowel. In addition, Elvin et al. (2021) reported that the categorization patterns of Portuguese Brazilian vowels by European Spanish listeners were largely predicted by the LDA model.

This study explores how English vowels are classified in terms of the phonetic categories of Cypriot Greek, a Greek variety with a simple vowel system. There are five pure vowel qualities, namely /i e a o u/, without any length distinctions, though stressed vowels can be longer than unstressed vowels (see Georgiou and Themistocleous 2021). According to Arvaniti (2010), Cypriot Greek vowels can phonetically be transcribed as [i ε a ɔ u]. Standard Modern Greek shares the same

qualities as Cypriot Greek, but there are some small acoustic differences. For example, Standard Modern Greek /e a o/ vowels are higher than the corresponding Cypriot Greek vowels, and the Standard Modern Greek unstressed /i a u/ vowels are more raised than the corresponding Cypriot Greek vowels (see Themistocleous 2017). Standard Southern British English (henceforth, SSBE) has a larger and more complex vowel system consisting of lax /ɪ ʊ e ə ʌ ɒ/ and tense /i: u: ɜ: ɔ: ɑ:/ (see Deterding 1997). The perception of English vowels by native speakers of Greek has been investigated in only a few previous studies. For example, Georgiou (2019a) examined the perception of English vowels by Cypriot Greek children with two different proficiency levels of English, finding that English /i i: e ə ʌ ɑ: ɒ ɔ: ʊ u:/ vowels were assimilated to Cypriot Greek /i i e a a o o u u/ respectively by both the low and the high proficiency children. Only the categorization of /ɜ:/ differed between the two groups, as it was assimilated to L1 /e/ by the former group, while it comprised a non-assimilated sound for the latter. It was observed that two or more L1 vowels were categorized into a single English category, demonstrating difficulty in the discrimination of some pairs of English vowels. Such difficulties are expected, since the two languages differ in the size and complexity of vowel system and there are significant acoustic differences between their sounds. Lengeris (2009) examined the perceptual assimilation of English vowels by Standard Modern Greek and Japanese speakers. The author concluded that English /i i: e ə ʌ ɑ: ɒ ɔ: ʊ u:/ were assimilated to Greek /i i e a a o o o u u/ respectively, similar to Georgiou (2019a). A more recent study examined the perception and production of English vowels /ɪ/ and /i:/ by Cypriot Greek adult speakers of L2 English (Georgiou 2022b), showing that both vowels were classified as Cypriot Greek /i/, supporting the findings of previous studies.

The purpose of this study is to investigate the classification of English vowels in terms of the phonetic categories of Cypriot Greek and the role of acoustic similarity in the prediction of the speakers' L2 perceptual patterns. This classification will be initially predicted through two trained LDA models, one including *F1* and *F2* values and another including *F1*, *F2*, and duration values of both Cypriot Greek and English vowels. The predictive model will be followed by a classification test in which Cypriot Greek speakers will be asked to classify L2 English vowels in terms of their L1 vowel categories. This will allow us to assess the role of acoustic similarity in making empirical predictions about the classification of L2 vowels using an LDA paradigm. Testing this role using different sets of languages (including under-researched ones such as Cypriot Greek) will help us to better define the contribution of acoustic similarity to L2 speech perception and inform the theories which suggest a link between the acoustic distance between L1 and L2 vowels and the perception of the L2 vowels. Importantly, the predictions are based on methods (i.e., chance scores) employed by a new speech model, the Universal Perceptual Model (UPM), which, like other popular models, aims to account for the difficulties of L2 speakers in the discrimination of L2 sound contrasts (see Georgiou 2021a, 2022b). UPM argues that the degree of overlap between non-native phonetic categories defines their perception. In addition, this study will allow the direct comparison of the accuracy of the two LDA models, offering important conclusions about the

significance of acoustic cues (either spectral or both spectral and durational) for the prediction of L2 sound classification. LDA models in speech perception studies are usually fed formant frequency values and duration (e.g., Kim and Clayards 2019) and rarely formant frequencies only (e.g., Li et al. 2021). To the best of our knowledge, no studies have compared LDA models trained on both formants and duration values with models trained on formants only. This study will therefore answer three main research questions: i) how do Cypriot Greek listeners of L2 English classify the English vowels in terms of their L1 vowel system, ii) to what extent does acoustic similarity of L1-L2 vowels predicts L2 vowel classification, and iii) how do LDA models with different input features differ with respect to the prediction of the classification of English vowels?

2. PRODUCTION STUDY

The production study served as the basis for the acoustic comparison of Cypriot Greek and English vowels and the development of the predictions regarding the classification of L2 vowels in terms of the speakers' L1 categories. This has been achieved with the use of an LDA paradigm.

2.1 Methodology

This section describes the participants, the stimuli, and the procedure of the study. The latter encompasses information about the analysis of the speech material and the use of the machine learning algorithm for the formulation of predictions.

2.1.1 Participants

Twenty-one participants were recruited for the purpose of the production study. Eleven individuals were native speakers of Cypriot Greek with an age range of 20-45 ($M_{\text{age}} = 32.2$, $SD = 7.71$). They were permanent residents of Cyprus and had a moderate socioeconomic status. They stated that they had never lived in an English-speaking country for more than one month. Ten L1 English (SSBE) speakers were also recruited. Their age ranged from 25-53 ($M_{\text{age}} = 39.8$, $SD = 9.43$) and they were permanent residents of Cyprus. They used their L1 during everyday communication and had no knowledge of Cypriot Greek. None of the participants had experienced any language, hearing, or cognitive disorder during their lifetime. All participants were female speakers to eliminate any gender bias in the results (see Yang 1992).

2.1.2 Stimuli

The five Cypriot Greek vowels /i e a o u/ comprised the stimuli of the study. The vowels were embedded in monosyllabic words with a /pVs/ context (V = vowel) and were part of the carrier phrase "Léne <target word> tóra," 'they say <target word> now'. The 11 English vowels, that is, lax /ɪ ʊ e æ ʌ ɒ/ and tense /i: u: ɜ: ɔ: ɑ:/ were embedded in monosyllabic words with a /hVd/ context and were part of the carrier phrase "They say <target word> now".

2.1.3 Procedure

All speakers completed the production task individually in quiet rooms. A list of the carrier phrases was presented to them, and they were asked to produce them as if speaking to a friend. The phrases were presented in the standard orthography of the speakers' L1s and their productions were recorded using a professional audio recorder at a 44.1 sampling rate. The phrases were randomized for each participant. Cypriot Greek speakers produced a total number of 220 items (5 vowels \times 4 repetitions \times 11 speakers). The same number of items was also produced by the English speakers (11 vowels \times 2 repetitions \times 10 speakers).

Speech analysis

Cypriot Greek speakers' output was manipulated using Audacity software and the target words were extracted and analyzed in Praat (Boersma and Weenink 2022). The boundaries of each vowel were defined through the visual analysis of spectrograms and waveforms to extract *F1*, *F2*, and duration. The following configurations were used: window length: 0.025 ms, pre-emphasis: 50 Hz, and spectrogram view range: 5500 Hz. The initial point of vowels' acoustic analysis was considered as the end of the quasi-periodicity of the preceding consonant /p/ for Cypriot Greek and /h/ for English and the onset point of vowel (V). The end point of the vowel was considered as the end of the quasi-periodicity of vowel (V) and the onset point of the second consonant /s/ for Cypriot Greek and /d/ for English. Formants were measured at their midpoint. The extraction of vowel duration was done through manual labelling of the starting and end points of each vowel token by the author, where the duration of the vowels was measured as the interval between the starting and ending point of the vocalic portion. *F1* and *F2* were normalized with the *Lobanov* method through the *vowels* package (Kendall and Thomas 2018). The normalized values were transformed into Hz.

Linear discriminant analysis (LDA)

The classification of English vowels in terms of Cypriot Greek phonetic categories was investigated using LDA. The analysis was conducted in R (R Core 2022) with the *MASS* package (Ripley et al. 2022), following the procedure used by Strange et al. (2005) and Gilichinskaya and Strange (2010). Two models were trained: one including only the first two formant frequencies (*F1*, *F2*) (henceforth 'F') and another including the two formant frequencies plus the duration of vowels (henceforth 'FD').

2.2 Results

The results of the production study show that English /i:/ is a close acoustic instance of Cypriot Greek /i/, while English /ɪ/ is between Cypriot Greek /i/ and /e/. English /e/ seems to spectrally overlap with Cypriot Greek /e/, while English /ɜ:/ is found between Cypriot Greek /e/ and /a/. In addition, English /æ/ and /ʌ/ are close acoustic exemplars of Cypriot Greek /a/. English /ɑ:/ is between Cypriot Greek /a/ and /o/, whereas English /ɒ/ overlaps Cypriot Greek /o/. Finally, English /ɔ:/ spectrally

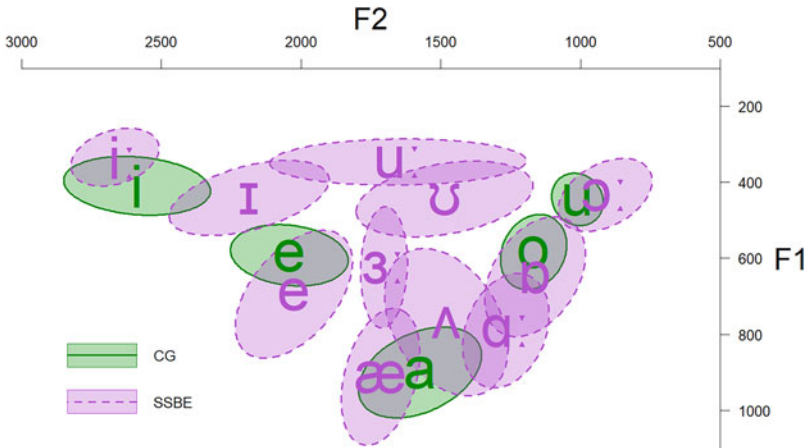


Fig. 1 $F1 \times F2$ (Hz) of Cyriot Greek (CG) and English (SSBE) vowels

overlaps Cyriot Greek /u/, while English /u:/ and /ʊ/ are close to several Cyriot Greek vowels, namely, /e/, /o/, and /u/. Figure 1 illustrates the $F1$ and $F2$ of Cyriot Greek and English vowels.

The results also indicate the duration of Cyriot Greek and English vowels. Cyriot Greek /a/ has the longest duration, while /i/ has the shortest one. The durational differences between Cyriot Greek vowels are slight. English vowels have varied durations. Lax vowels have considerably lower durations than tense vowels. English lax vowels approach Cyriot Greek vowels in terms of duration to a greater extent compared to tense vowels.

The cross-validation method was used to estimate the accuracy of the LDA predictive models in practice. It was found that the FD model demonstrated 97.7%

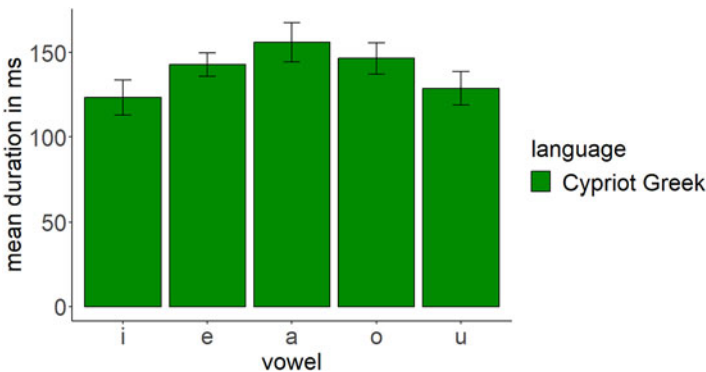


Fig. 2 Vowel duration and SDs of Cyriot Greek vowels

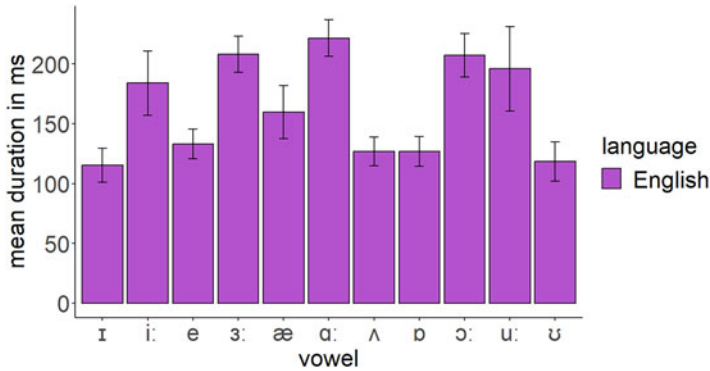


Fig. 3 Vowel duration and *SDs* of English vowels

correct classification accuracy, while the F model demonstrated 96.3% correct classification accuracy. In other words, both models yielded high classification accuracy. It was therefore possible to supply the trained models with the measurements of English vowels to predict how English vowels would be classified in terms of Cypriot Greek categories. Table 1 shows the proportions of classification of the English vowels in terms of the speakers' L1 categories.

In the FD model, English /ɪ/ was classified with the highest proportion in terms of /i/, while in the F model it was classified in terms of /e/. Also, in both models, English /i:/, /e/, and /æ/ were classified in terms of Cypriot Greek /i/, /e/, and /a/ respectively. English /ɜ:/ showed different patterns in the two models as it was classified with the highest proportion in terms of Cypriot Greek /a/ in the FD model, but as Cypriot Greek /e/ in the F model. A different pattern was also indicated for English /ɑ:/ since it was classified in terms of Cypriot Greek /a/ in the FD model, while it was mostly classified as a response of Cypriot Greek /o/ in the F model. English /ʌ/ was classified with the highest proportion in terms of Cypriot Greek /o/ in the FD model but in terms of /a/ in the F model. English /ɒ/ was classified in terms of Cypriot Greek /o/ in both models. Furthermore, English /ɔ:/ was a response to Cypriot Greek /o/ in the FD model, but Cypriot Greek /u/ in the F model. The classification of English /ʊ/ was similar in the two models as it was classified with the highest proportion as Cypriot Greek /u/. Finally, English /u:/ was classified with the highest proportion as Cypriot Greek /e/ in both models. Thus, LDA indicates similarities but also several differences between the two models regarding the classification of English vowels.

The acoustic classification results can be compared to the perceptual assimilation of English vowels by Cypriot Greek speakers. Specifically, English vowels /ɪ i: e æ ɑ: ʊ ɒ ɔ:/ in the FD model were classified with the highest proportion in terms of the same L1 categories as those indicated by Georgiou (2019a), while this was the case for English vowels /i: e ɜ: ʌ æ ʊ ɒ/ in the F model. It is evident that the classification of some vowels (e.g., /ɪ ɑ: ɔ:/) as predicted by the FD model of LDA are closer to the results reported by Georgiou (2019a) in comparison to the corresponding

vowels		FD					F				
		Cypriot Greek									
		i	e	a	o	u	i	e	a	o	u
English	ɪ	0.60	0.40				0.45	0.55			
	i:	1.00					1.00				
	e		1.00					0.95	0.05		
	ɜ:		0.40	0.60				0.85	0.15		
	æ		0.05	0.95				0.05	0.95		
	ɑ:			0.95	0.05				0.45	0.55	
	ʌ		0.10	0.40	0.50			0.10	0.55	0.35	
	ɒ			0.05	0.65	0.30			0.05	0.90	0.05
	ɔ:				1.00					0.10	0.90
	ʊ		0.30		0.15	0.55			0.30	0.20	0.50
u:	0.05	0.50		0.45		0.05	0.50			0.45	

Table 1: Classification results of the LDA analyses for the FD and F models. Shaded cells represent L1 responses with the highest proportion. Bold represents above chance responses (≥ 0.20). Chance score is calculated by dividing 100 (or 1.00) by the number of the script responses; here 5 ($1.00/5 = 0.20$)

vowels predicted by the F model. However, some vowels (e.g., /ɜ: ʌ/) as predicted by the F model are closer to those indicated by the previous study compared to the corresponding vowels predicted by the FD model.

3. PERCEPTUAL STUDY

The perceptual study examined the classification of L2 vowels in terms of the speakers' L1 vowels, aiming to verify the predictions of the production study. For this purpose, a classification test was used.

3.1 Methodology

Details regarding the participants, stimuli, and procedural aspects of the perceptual study are outlined below.

3.1.1 Participants

Twelve Cypriot Greek speakers ($n_{\text{females}} = 6$) took part in the perception study. Their age range was 21-42 ($M_{\text{age}} = 31.92$, $SD = 5.98$) and they were permanent residents of Cyprus. According to their reports, they had a mean age of English learning onset of 8 years ($SD = 1.58$). The mean hours of daily use of English were 2.58 ($SD = 2.06$), while the mean hours of daily input in English were 2.83 ($SD = 1.48$). All reported good knowledge of English at the B2/C1 levels and their self-perceived score of English understanding skills was 4.67/5 ($SD = 0.47$). Despite the participants' advanced knowledge of English, it is expected that their L1 will have a great influence on the perception of L2 vowels (see also Georgiou 2022b) due to the fact that they learnt the language academically rather than naturalistically and that pronunciation teaching is often marginalized in English language classrooms in Cyprus (Georgiou 2019b). All participants had healthy hearing and were free from any language or cognitive problems.

3.1.2 Stimuli

The stimuli consisted of the 11 English monophthongs embedded in a /hVd/ word context. These words were part of the carrier phrase "They say <target word> now", as in the production study. Two female adult English (SSBE) speakers were asked to produce these phrases naturally and their productions were recorded using a professional audio recorder at a 44.1 kHz sampling rate. The output was normalized for peak intensity in Praat. The $F1 \times F2$ of their productions are illustrated in Figure 4.

3.1.3 Procedure

The participants were tested individually in quiet rooms. The test was prepared in a Praat script, which was presented to the participants through a PC monitor. Participants were instructed to take a seat in front of the monitor and follow the instructions. They listened to the words including the target English vowels and

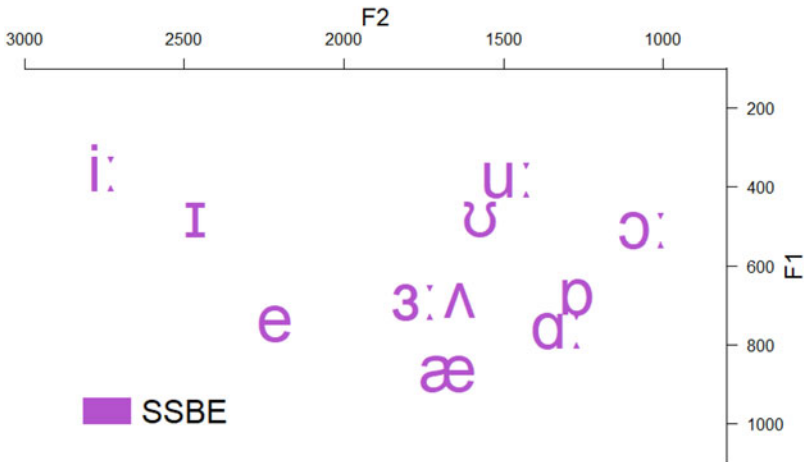


Fig. 4 $F1 \times F2$ of English (SSBE) vowels produced by the two English speakers

had to click on the script label that was acoustically the most similar exemplar to the vowel they heard. The labels included the orthographic representation of the five Cypriot Greek vowels, namely, ‘ι’, ‘ε’, ‘α’, ‘ο’, and ‘ου’. They were also asked to rate how “good” or “bad” an exemplar that English vowel was by selecting a rating from one (very poor) to five (very good). The speakers classified a total number of 48 trials (11 vowels \times 4 repetitions). The interval between a click and the presentation of the next trial was 500 ms. Although there was no time restriction, the participants were asked to provide rapid responses to the script. No feedback was given and the acoustic stimuli could not be repeated. Prior to the main experiment, the participants completed a familiarization task with 4 test items. The classification test lasted about 20-25 minutes for each participant, with an optional five-minute break at the midpoint.

3.2 Results

Classification data were based on the framework of the UPM model. In this respect, English /i i:/ were classified as above chance responses (i.e., responses of which the classification proportion exceeded the chance score from it) as Cypriot Greek /i/. English /e ɜ:/ were classified as above chance responses as Cypriot Greek /e/, English /æ/ as English /a/, and English /ɑ: ʌ/ as Cypriot Greek /a/ and /o/ respectively. English /ɒ/ was classified as Cypriot Greek /o/, while English /ɔ: ɒ:/ was classified as an above chance response in both Cypriot Greek /o/ and /u/. Finally, English /ʊ u:/ were both classified as Cypriot Greek /u/. All responses but /ɑ: ʌ ɒ ɔ:/ were optimal to the respective L1 categories (i.e., the only above chance responses to that L1 category). [Table 2](#) presents the classification of English vowels in terms of Cypriot Greek categories.

These results were compared with those presented by the LDA analyses. The FD model accurately predicted the classification of English /i i: e æ ɑ: ɒ ʊ/ as responses

vowels		Cypriot Greek				
		i	e	a	o	u
English	i	0.85	0.15			
	i:	0.98	0.02			
	e		1.00			
	ɜ:		0.90	0.10		
	æ	0.02		0.92	0.06	
	ɑ:	0.08		0.65	0.27	
	ʌ	0.04		0.75	0.21	
	ɒ	0.06		0.29	0.63	0.02
	ɔ:				0.50	0.50
	ʊ					1.00
u:					1.00	

Table 2: Classification of L2 English vowels in terms of Cypriot Greek categories. Shaded cells represent L1 responses with the highest proportion. Bold represents above chance responses (≥ 0.20).

with the highest proportion in terms of a single L1 category. The F model successfully predicted the classification of English /i: e ɜ: æ ʌ ɒ ʊ/. Both models failed to predict the classification of English /u:/ and /ɔ:/, showing that the former was classified as a response with the highest proportion to Cypriot Greek /e/ rather than /u/ and that the latter was classified as a response of Cypriot Greek /o/ (FD) and /u/ (F) rather than as a response to both Cypriot Greek /o/ and /u/. With regard to the predictions of the full ranges of responses exceeding the chance score, the FD model successfully predicted the classification of English /i: e æ/, while the F model successfully predicted the classification of English /i: ɜ: e æ ʌ/.

4. DISCUSSION

This article investigated the classification of L2 English vowels in terms of L1 Cypriot Greek vowel categories. It also aimed to examine whether L1-L2 acoustic similarity can predict the classification of L2 vowels. The classification predictions were based on two LDA models with different input features (formants and duration versus duration only). Native Cypriot Greek speakers performed a perceptual task to investigate how they classify L2 English vowels as categories of their L1 and how successful the consideration of L1-L2 acoustic similarity is for the prediction of vowel classification patterns.

The results of the perceptual task showed that most L2 vowels were optimal responses to an L1 category. This indicates that speakers realize most L2 vowels as being acoustically close to a particular L1 vowel. For example, the almost perfect classification of English /i:/ in terms of Cypriot Greek /i/ is predictable since this L2 vowel is the closest acoustic exemplar of Cypriot Greek /i/. While

most of the L2 vowels were optimal responses of an L1 category, English /ɔ:/ was classified in terms of two L1 vowel categories. The classification of English /ɔ:/ vowel in a range of L1 categories can be explained by the fact that it shares common features with different L1 vowels (e.g., spectrally it overlaps Cypriot Greek /u/, but temporally it is closer to Cypriot Greek /o/). More than one English vowel was classified as an optimal response in terms of the same Cypriot Greek vowel category. Specifically, English /i i:/ were classified as Cypriot Greek /i/, English /e ɜ:/ as Cypriot Greek /e/, and English /u u:/ as Cypriot Greek /u/. Apart from the acoustic-phonetic differences between L1 and L2 vowels, specifically the close acoustic distance of an L1 vowel with several L2 vowels, this might also be favoured by the small vowel inventory of Cypriot Greek, which “compels” speakers to find the most acoustically approximate corresponding vowel to a particular L2 vowel from a narrow selection of L1 responses. It is obvious that every L2 sound is perceived through the lens of the speakers’ L1 and that there is a common acoustic space that accommodates both L1 and L2 sounds (see Flege et al. 2003). The findings of this study can be compared with previous findings provided by Georgiou (2019a), who investigated the assimilation of English vowels by Cypriot Greek child learners of English. When the L1 responses with the highest proportion of classification are considered, similar mappings are observed in the two studies. Specifically, English /i i: e ʌ æ α: ɒ u u:/ were mostly classified in terms of Cypriot Greek /i i e a a o u u/ in both studies. Only the classification of /ɔ:/ differed; in the present study, it was classified equally in terms of both L1 /a/ and /o/, while in the previous study it was highly assimilated to L1 /o/.

Perceptual mapping of L2 sounds to L1 sound categories may predict the discrimination of L2 sound contrasts (see Best and Tyler 2007, Georgiou et al. 2020b). UPM distinguishes between completely overlapping contrasts (those sharing the same above-chance L1 category or set of L1 categories), partially overlapping contrasts (those sharing at least one above-chance category), and nonoverlapping contrasts (those sharing no above-chance categories). In the lens of UPM predictions, /i – i:/, /e – ɜ:/, /a: – ʌ/, /ɒ – ʌ/, and /u – u:/ would be completely overlapping contrasts, yielding poor discrimination. Contrasts such as /ɒ – ɔ:/, /u – ɔ:/, and /u: – ɔ:/ /æ – α:/ would be considered partially overlapping contrasts and therefore their discrimination would be better than those of completely overlapping contrasts. Other contrasts such as /i – e/, /u: – æ/, /ʌ – ɜ:/, etc. would be nonoverlapping and thus their discrimination would be better than that of the other two types of contrasts. The purpose of this study was not to assess the discrimination accuracy of L2 contrasts, but to investigate the speakers’ L2 vowel classification patterns and the role of acoustic similarity in speech perception. A future study may examine speakers’ discrimination patterns to test the assumptions underlying UPM and its ability to accurately predict the discrimination accuracy from the classification of L2 speech sounds in terms of L1 categories.

Another question that this study aimed to answer was the capacity of L1-L2 vowel acoustic similarity to predict speakers’ L2 vowel classification patterns. Both LDA models were successful in the prediction of 7 out of 11 English vowels as responses with the highest proportion for a single L1 category. Therefore, L1-

L2 sound acoustic similarity can be a good indicator of sound classification, as the highest classification proportion for the majority of English vowels was successfully predicted with the use of LDA models. However, each model could not predict the classification of 4 out of 11 English vowels. Similarly, Gilichinskaya and Strange (2010) observed that although acoustic similarity successfully predicted the categorization of American English vowels into L1 Russian categories, it failed to predict the mapping of American English /ε/. Elvin et al. (2021) found that acoustic similarity successfully indicated the categorization of Brazilian Portuguese vowels by speakers of Australian English and European Spanish. However, the authors found that some categorization responses from Australian English listeners could not be predicted, and that perceptual similarity was a better predictor of the discrimination accuracy of European Spanish listeners than acoustic similarity. The discrepancies found in previous studies and this study may be attributed to the absence of important input features in LDA such as *F3*, which defines lip rounding and the lengthening of the vocal tract, and dynamic formant trajectories, which comprise important aspects of vowel perception and production (Elvin et al. 2016, Escudero et al. 2018, Williams et al. 2018).

Both LDA models correctly predicted the highest classification proportion for most L2 vowels. They could not predict 4 out of 11 vowels; two of these were the same in both models. These models can provide useful information on the cues needed to develop accurate predictions about the classification of L2 vowels since the FD model uses formant frequencies and duration as input, while the F model uses formant frequencies only. FD, but not F, predicted with success English /ɪ ɑ:/, indicating that both the spectral and the duration features of these vowels were significant for the prediction of speakers' classification patterns. In addition, F, but not FD, predicted with success English /ɜ: ʌ/, signaling that the spectral characteristics of these vowels were most important for the classification predictions. Thus, it can be argued that the importance of cues may depend on the vowel. In addition, in terms of accurate prediction (e.g., predicting the whole range of responses exceeding the chance score), the FD model predicted three vowels (/i: e æ/) and the F model predicted five vowels (/i: e ɜ: æ ʌ/). The F model was therefore more successful than the FD model in predicting the full range of responses. This indicates that duration cues might be less important than formant cues for predicting the perception of L2 English vowels by Cypriot Greek speakers. Although this study did not directly examine the reliance of speakers on acoustic cues, the above conclusions may be related to earlier findings suggesting that Cypriot Greek speakers prioritize the use of spectral cues during the categorization of English L2 vowels and use temporal cues when access to spectral information is limited (Georgiou 2019a).

5. CONCLUSIONS

The effect of the speakers' L1 on their categorization of L2 vowels is reflected in the fact that several L2 vowels highly overlapped with a single L1 vowel category, which can potentially create difficulty in the discrimination of particular L2 sound contrasts.

In addition, the results showed that crosslinguistic acoustic similarity can be quite successful in predicting the classification of L2 vowels. However, more accurate predictions could possibly be achieved by feeding the LDA models more input parameters. In addition to the theoretical contribution of these findings, they may also have pedagogical value. For example, a tentative description of the acoustic differences between L1 and L2 sounds can provide the opportunity to predict students' perceptual (and production) difficulties and guide the teaching of L2 speech sounds in a particular direction. Finally, this study aimed to examine the classification of L2 vowels in terms of the speakers' L1 categories and the ability of acoustic similarity to predict this classification. A future study may compare the accuracy of acoustic versus perceptual similarity on the basis of a discrimination test in which Cypriot Greek speakers will discriminate L2 English sound contrasts. The theoretical predictions of the UPM model can be used for this purpose.

REFERENCES

- Arvaniti, Amalia. 2010. Linguistic practices in Cyprus and the emergence of Cypriot Standard Greek. *Mediterranean Language Review* 17: 15–45.
- Baigorri, Miriam, Luca Campanelli, and Erika S. Levy. 2019. Perception of American–English vowels by early and late Spanish–English bilinguals. *Language and speech* 62(4): 681–700.
- Best, Catherine T. 1995. A direct realist view of cross-language speech perception: New directions in research and theory. In *Speech perception and linguistic experience: Theoretical and methodological issues*, ed. Winifred Strange, 171–204. Baltimore: York Press.
- Best, C. T., & Tyler, M. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In O-S. Bohn & M. J. Munro (Eds.), *Second language speech learning: In honor of James Emil Flege* (pp. 13-34). John Benjamins.
- Boersma, Paul and David Weenink. 2022. *Praat: doing phonetics by computer* [Computer program]. <http://www.fon.hum.uva.nl/praat/>
- Darcy, Isabelle, and Franziska Krüger. 2012. Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics* 40(4): 568–581.
- Deterding, David. 1997. The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association* 27(1-2): 47–55.
- Elvin, J., Williams, D., and Escudero, P. (2016). Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English. *The Journal of the acoustical society of America*, 140(1), 576-581.
- Elvin, Jaydene, Paola Escudero, and Polina Vasiliev. 2014. Spanish is better than English for discriminating Portuguese vowels: Acoustic similarity versus vowel inventory size. *Frontiers in Psychology* 5: 1188.
- Elvin, Jaydene, Daniel Williams, Jason A. Shaw, Catherine T. Best, and Paola Escudero. 2021. The role of acoustic similarity and non-native categorisation in predicting non-native discrimination: Brazilian Portuguese vowels by English vs. Spanish listeners. *Languages* 6(1): 44.
- Escudero, Paola. 2009. Linguistic perception of similar L2 sounds. In *Phonology in perception*, ed. Paul Boersma and Silke Hamann, 151–190. Berlin: Mouton de Gruyter.
- Escudero, Paola, Karen E. Mulak, Jaydene Elvin, and Nicole M. Traynor. 2018. Mummy, keep it steady: Phonetic variation shapes word learning at 15 and 17 months. *Developmental Science* 21(5): e12640.

- Flege, James Emil. 1995. Second language speech learning: Theory, findings and problems. In *Speech perception and linguistic experience: Theoretical and methodological issues*, ed. Winifred Strange, 233–277. Baltimore: York Press.
- Flege, James Emil. 2003. Assessing constraints on second-language segmental production and perception. In *Phonetics and phonology in language comprehension and production: Differences and similarities*, ed. Niels O. Schiller and Antje S. Meyer, 319–355. Berlin: Mouton De Gruyter.
- Flege, James Emil., and Ocke-Schwen Bohn. 2021. The Revised Speech Learning Model (SLM-r). In *Second language speech learning: Theoretical and empirical progress*, ed. Ratreë Wayland, 3–83. Cambridge: Cambridge University Press.
- Flege, James E., Carlo Schirru, and Ian R.A. MacKay. 2003. Interaction between the native and second language phonetic subsystems. *Speech Communication* 40(4): 467–491.
- Fox, Robert Allen, James Emil Flege, and Murray J. Munro. 1995. The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *The Journal of the Acoustical Society of America* 97(4): 2540–2551.
- Georgiou, Georgios P. 2019a. Bit and beat are heard as the same: Mapping the vowel perceptual patterns of Greek-English bilingual children. *Language Sciences* 72: 1–12.
- Georgiou, Georgios P. 2019b. EFL teachers' cognitions about pronunciation in Cyprus. *Journal of Multilingual and Multicultural Development* 40(6): 538–550.
- Georgiou, Georgios P. 2021a. Toward a new model for speech perception: The Universal Perceptual Model (UPM) of second language. *Cognitive Processing* 22: 277–289.
- Georgiou, Georgios P. 2021b. Interplay between perceived cross-linguistic similarity and L2 production: Analyzing the L2 vowel patterns of bilinguals. *Journal of Second Language Studies* 4(1): 48–64.
- Georgiou, Georgios P. 2022a. The impact of auditory perceptual training on the perception and production of English vowels by Cypriot Greek children and adults. *Language Learning and Development* 18(4): 379–392.
- Georgiou, Georgios P. 2022b. The acquisition of /i/–/i:/ is challenging: Perceptual and production evidence from Cypriot Greek speakers of English. *Behavioral Sciences* 12(12): 469.
- Georgiou, Georgios P., and Charalambos Themistocleous. 2021. Vowel learning in diglossic settings: Evidence from Arabic-Greek learners. *International Journal of Bilingualism* 25(1): 135–150.
- Georgiou, Georgios P., Natalia V. Perfilieva, and Maria Tenizi. 2020a. Vocabulary size leads to better attunement to L2 phonetic differences: Clues from Russian learners of English. *Language Learning and Development* 16(4): 382–398.
- Georgiou, Georgios P., Natalia V. Perfilieva, Vladimir N. Denisenko, and Natalia V. Novospasskaya. 2020b. Perceptual realization of Greek consonants by Russian monolingual speakers. *Speech Communication* 125: 7–14.
- Gilichinskaya, Yana D., and Winifred Strange. 2010. Perceptual assimilation of American English vowels by inexperienced Russian listeners. *The Journal of the Acoustical Society of America* 128(2): EL80–EL85.
- Grimaldi, Mirko, Bianca Sisinni, Barbara Gili Fivela, Sara Invitto, Donatella Resta, Paavo Alku, and Elvira Brattico. 2014. Assimilation of L2 vowels to L1 phonemes governs L2 learning in adulthood: A behavioral and ERP study. *Frontiers in Human Neuroscience* 8: 279.
- Hacquad, Valentine, Mary Ann Walter, and Alec Marantz. 2007. The effects of inventory on vowel perception in French and Spanish: An MEG study. *Brain and Language* 100(3): 295–300.

- Iverson, Paul, and Bronwen G. Evans. 2007. Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *The Journal of the Acoustical Society of America* 122(5): 2842–2854.
- Iverson, Paul, and Bronwen G. Evans. 2009. Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America* 126(2): 866–877.
- Iverson, Paul, Patricia K. Kuhl, Reiko Akahane-Yamada, Eugen Diesch, Andreas Kettermann, and Claudia Siebert. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87(1): B47–B57.
- Kartushina, Natalia, and Ulrich H. Frauenfelder. 2014. On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology* 5: 1246.
- Kendall, Tyler, and Eric R. Thomas. 2018. *Vowel Manipulation, Normalization, and Plotting*. R package version 1.2-2.
- Kim, Donghyun, and Meghan Clayards. 2019. Individual differences in the link between perception and production and the mechanisms of phonetic imitation. *Language, Cognition and Neuroscience* 34(6): 769–786.
- Klecka, William R. 1980. *Discriminant analysis: Quantitative applications in the social sciences*. Newbury Park: Sage.
- Lengeris, Angelos. 2009. Perceptual assimilation and L2 learning: Evidence from the perception of Southern British English vowels by native speakers of Greek and Japanese. *Phonetica* 66(3): 169–187.
- Li, Mingtao, Sio Hang Pun, and Fei Chen. 2021. A preliminary study of classifying spoken vowels with EEG signals. In *Proceedings of the 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, 13–16. DOI: [10.1109/NER49283.2021](https://doi.org/10.1109/NER49283.2021)
- Melnik-Leroy, Gerda Ana, Rory Turnbull, and Sharon Peperkamp. 2022. On the relationship between perception and production of L2 sounds: Evidence from Anglophones' processing of the French /u/–/y/ contrast. *Second Language Research* 38(3): 581–605.
- R Core Team. 2022. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Ripley, Brian, Bill Venables, Douglas M. Bates, Kurt Hornik, Albrecht Gebhardt, and David Firth. 2022. *Support functions and datasets for Venables and Ripley's MASS*. R package version 7.3-53.1.
- Souza, Hanna Kivistö-de, Angélica Carlet, Izabela Anna Jułkowska, and Anabela Rato. 2017. Vowel inventory size matters: Assessing cue-weighting in L2 vowel perception. *Ilha do Desterro* 70: 33–46.
- Strange, Winifred, Ocke-Schwen Bohn, Kanae Nishi, and Sonja A. Trent. 2005. Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. *The Journal of the Acoustical Society of America* 118(3): 1751–1762.
- Themistocleous, Charalambos. 2017. The nature of phonetic gradience across a dialect continuum: evidence from modern Greek vowels. *Phonetica* 74(3): 157–172.
- Williams, Daniel, Paola Escudero, and Adamantios Gafos. 2018. Spectral change and duration as cues in Australian English listeners' front vowel categorization. *The Journal of the Acoustical Society of America* 144(3): EL215–EL221.
- Yang, Byunggon. 1992. An acoustical study of Korean monophthongs produced by male and female speakers. *The Journal of the Acoustical Society of America* 91(4): 2280–2283.