Available online at www.sciencedirect.com

ScienceDirect

Lingua 310 (2024) 103805

Lingua

www.elsevier.com/locate/lingua

Greek-Cypriot learners' perception and production of L2 English vowels

Dimitra Dimitriou

University of Central Lancashire Cyprus, 12-14 University avenue, 7080 Pyla, Larnaca, Cyprus

Received 30 January 2024; revised 9 August 2024; accepted in revised form 19 August 2024;

Abstract

Acquiring L2 segments is particularly challenging for L2 learners, especially when the L1 and L2 inventories involve different contrasts and acoustic cues. The present research investigated the perception and production of L2 English vowels by adult Greek-Cypriot learners and examined the effects of orthographic cues and consonantal context in their performance. Perceptual performance was assessed through a forced-choice identification task and production performance through a wordlist-reading and an elicitation task, both analyzed acoustically and through intelligibility ratings. The findings showed the influence of the L1 on both the perception and production of L2 segments, supporting the assumptions of current models of speech perception and production. Learners faced challenges in perceiving the members of L2 contrasts and mostly used their L1 articulatory routines in their productions of L2 vowels. Orthographic cues or consonantal context did not significantly affect learners' productions or overall perception, although strong contextual effects were observed in individual target vowels. This study is the first to provide an in-depth comparison of CYG and English vowels as well as an examination of the acquisition of L2 English vowels by this population, which can guide EFL teachers and course developers in developing appropriate EFL curricula that incorporate pronunciation instruction.

Keywords: Vowel perception; Vowel production; Greek-Cypriot learners; L1-L2 interaction; Orthographic cues; Consonantal context

1. INTRODUCTION

Research in the field of Foreign Language Acquisition (FLA) demonstrates that learning to accurately perceive and produce new phonetic categories is one of the most challenging tasks L2 learners face (e.g. Dimitriou, 2019; Iverson et al., 2003; Iverson and Evans, 2007; Lengeris, 2009). Previous studies show that learners whose L1 has a smaller vowel inventory than the L2 inventory, such as Spanish or Greek learners of English, have difficulties in discriminating English vowel contrasts (Cebrian, 2006; Iverson & Evans, 2009; Lengeris, 2009; Sakai, 2016), since in such cases, a single category exists in the acoustic space occupied by two or more L2 vowels, meaning that the two unfamiliar L2 sounds are likely to be perceived as exemplars of the same category by learners, who are not accustomed to attending

E-mail address: ddimitriou4@uclan.ac.uk

https://doi.org/10.1016/j.lingua.2024.103805
0016-7037/© 2024 The Author(s). Published by Elsevier B.V.

to the appropriate acoustic cues to distinguish between members of a contrastive pair (e.g. Aliaga-Garcia & Mora, 2009; Cebrian, 2006; Cebrian, 2007; Flege, 1995; Flege & Bohn, 2021).

The present research investigates the perception and production of L2 English vowels by L1 CYG learners. CYG is a dialect of Standard Modern Greek (SMG) spoken by approximately 1 million speakers, mainly in Cyprus, but also by the immigrant communities in the UK, North America and Australia (Arvaniti, 1999; Simaki et al., 2015). CYG speakers learn English in an FLA context, mostly through formal instruction largely confined to the L2 classroom, which entails major limitations in terms of speech acquisition, since it lacks naturalistic exposure and provides little opportunity for learners to encounter native-like input (Barriuso & Hayes-Harb, 2018; Fabra & Romero, 2012; Georgiou, 2019; Hutchinson & Dmitrieva, 2022), especially since L2 pronunciation instruction is highly neglected (Kyprianou, 2015).

As a result of the general assumption that the variety of Greek used in Cyprus and the standard variety spoken in Greece are similar, the CYG dialect is generally neglected in the literature. However, examining CYG as a separate variety from SMG is important, since substantial variation can be found even in simple five-vowel systems as an effect of sociolinguistic factors and linguistic context (Arvaniti, 2010). Apart from subtle differences between the vowel inventories of SMG and CYG (Themistocleous, 2017a; Themistocleous, 2017b; Themistocleous & Logotheti, 2016), the political and historical background of Cyprus as a former British colony and the resulting colonial links between the countries may affect these learners differently, possibly providing additional motivation or affecting their language use patterns.

Despite this, previous research on the phonetics and phonology of CYG is limited and rarely focuses on vowels, while previous studies examining L2 English vowel acquisition by SMG or CYG learners have focused on the perception modality, paying little attention to production patterns (e.g. Georgiou, 2019; Lengeris, 2009; Lengeris & Hazan, 2007). Although the vowels of the dialect have been acoustically analyzed and compared to SMG vowels (Themistocleous, 2017a; Themistocleous, 2017b; Themistocleous & Logotheti, 2016), no previous study has compared CYG to English vowels to determine differences in their acoustic space, which may pose considerable problems for CYG learners.

More specifically, the CYG vowel system consists of five well-separated vowels (/i e a o u/) and does not have tense-lax or short-long distinctions (Arvaniti, 1999; Coutsougera, 2007; Georgiou, 2019; Themistocleous, 2017a; Themistocleous, 2017b; Themistocleous & Logotheti, 2016). On the other hand, the vowel system of English is larger, more complex, and exhibits regional variation (Hughes et al., 2013). Standard Southern British English (SSBE), in particular, contains five tense (/iː, uː, ɜː, ɔː, ɑː/) and six lax (/ɪ, ʊ, e, æ, ʌ, ɒ/) vowels (Bohn & Steinlen, 2003; Deterding, 1997). Although tense vowels are typically phonetically longer than their lax counterparts (Leung et al., 2016), the primary acoustic cue used to differentiate English tense and lax vowels is spectral quality, with duration being a secondary cue used in some varieties (Rato & Carlet, 2020).

2. LITERATURE REVIEW

Current models of L2 speech perception and production argue that the relationship between the L1 and L2 sound inventories may enable predictions as to the ease or difficulty with which learners will acquire an L2 sound. For instance, according to the Speech Learning Model (SLM; Flege, 1995) and its recent revision, SLM-r (Flege & Bohn, 2021), and the Perceptual Assimilation Model (PAM; Best, 1995) and PAM-L2 (Best & Tyler, 2007), which are the dominant models accounting for the formation of new categories by learners, the L1 has a significant influence over L2 phonological acquisition, and difficulties in the perception and production of L2 segments are at least to some extent predictable from the L1 and the acoustic similarity or dissimilarity between the L1 and L2 phonemes.

More specifically, the SLM (Flege, 1995) and the SLM-r (Flege & Bohn, 2021) posit that L1 and L2 phonetic categories exist in a common phonological space and interact through the processes of phonetic category assimilation and phonetic category dissimilation. Phonetic category assimilation occurs when the creation of a new category is blocked due to equivalence classification, i.e. the perception of an L2 sound as phonetically similar to an L1 sound, leading to the production of L2 sounds using the articulatory routines of L1 sounds (Flege, 1995). Phonetic category dissimilation, on the other hand, occurs when a new phonetic category is ultimately formed for an L2 sound (Flege, 1995). According to these assumptions, the establishment of a new category is more likely when the L2 sound is perceived to be more distant from the closest L1 sound (Flege, 1995). The SLM also assumed a unidirectional relationship between perception and production, whereby the former preceded the latter; however, as a result of contradictory findings in subsequent research, this assumption has been revised in the SLM-r, which argues for a bidirectional link whereby “segmental production and perception coevolve without precedence” (Flege & Bohn, 2021, p. 64).

Indeed, the nature of the interaction between the two modalities has been the subject of investigation of various previous studies, although a consensus has yet to be reached. For example, many studies report at least a modest relationship between the two modalities (e.g. Baker & Trofimovich, 2006; Kluge et al., 2007; Levy & Law, 2010; Melnik-Leroy et al., 2022; Zhang & Peng, 2017), while others found no correlation between them at all (e.g. Kartushina & Frauenfelder, 2014; Peperkamp & Bouchon, 2011). The direction of the link is not yet established either. The initial hypothesis of the SLM that accurate perception precedes accurate production is supported by some experimental studies (e.g. Casillas, 2019; Melnik-Leroy et al., 2022; Nagle, 2018), but others have found accurate production despite an inaccurate perception of an L2 contrast (e.g. Bohn & Flege, 1997) or mixed results such as better perception for some sounds and better production for others (e.g. Hao & de Jong, 2016).

The PAM (Best, 1995) and the PAM-L2 (Best & Tyler, 2007) focus on the perception of non-native phonemes, providing category assimilation patterns across languages. According to these models, when a new sound is encountered, learners attempt to assimilate it to perceptual categories already in place for the L1. If an L2 phoneme is perceived to be very similar to an L1 phoneme, then it is more likely that it will be assimilated to the L1 category (Best, 1995; Best & Tyler, 2007). According to these models, there are six patterns of assimilation of non-native phonological contrasts onto native categories: Two-Category Assimilation (TC type), Category-Goodness Difference (CG type), Single-Category Assimilation (SC type), Uncategorised-Uncategorised (UU type), Uncategorised-Categorised (UC type), and Non-assimilated (NA type).

These can in turn enable predictions as to how accurately the contrasts will be discriminated. For instance, TC contrasts are predicted to be easier for learners, followed by CG contrasts, where L2 phonemes are assimilated to a single L1 phonological category but with a different goodness of fit. On the other hand, SC contrasts, which involve two sounds being assimilated to a single L1 category as equally good or poor exemplars of it, are predicted to be particularly challenging for learners. Finally, the discrimination of UC and NA contrasts is predicted to be very good, whereas the discrimination of UU contrasts, in which neither of the L2 sounds falls within an L1 category, may vary from poor to very good, depending on their similarity to an L1 category (Best, 1995; Tyler, 2019). While these hypotheses are made for learners in immersion contexts, Tyler (2019) notes that the predictions of the PAM-L2 remain the same for FLA contexts, although CG and SC assimilations are less likely to be acquired in the EFL classroom.

Previous studies on adult SMG (Lengeris, 2009) and child CYG (Georgiou, 2019) learners of English found that the English vowels /iː, ɪ/, /e, ɜː/, /æ, ʌ/, /ɒ, ɔː/ and /ʊ, uː/ are mostly assimilated to Greek /i/, /e/, /a/, /o/ and /u/, respectively. One difference is that in Georgiou (2019), the /ɑː/ vowel was assimilated to the CYG category for /a/ instead of /o/ as Lengeris (2009) found, possibly due to differences in age, experience or native variety between the two groups. Similar results have been reported for Catalan and Spanish learners of English, due to the lack of a tense-lax distinction in the L1, which causes the assimilation of the English vowels to a single L1 category (Aliaga-Garcia & Mora, 2009; Carlet & Cebrian, 2014; Cebrian, 2019; Cebrian et al., 2019).

Importantly, the perceptual patterns reported in Georgiou (2019) can enable predictions as to the production of English vowels as well, since these learners are expected to use their five L1 vowels for the production of the L2 vowels similarly to their perceptual patterns. Georgiou (2019) reports a CG or UC assimilation pattern for the English vowel contrasts /ɪ/-/iː/, /e/-/ɜː/ and /ʊ/-/uː/, and an SC assimilation pattern for the English vowel contrasts /æ/-/ʌ/ and /ɒ/-/ɔː/. Assuming that better discrimination will lead to better production, and based on the PAM's prediction that SC contrasts are the most difficult to discriminate, it is expected that CYG learners of English will find it more difficult to accurately produce the vowels /æ/-/ʌ/ and /ɒ/-/ɔː/.

In addition to phonemic inventory differences, the two languages also differ in orthographic patterns, which may influence the way a word is perceived and produced (e.g. Bassetti, 2017; Bassetti et al., 2015; Nimz & Khattab, 2020; Stoehr & Martin, 2022). More specifically, as opposed to English, the SMG orthography is transparent, and although CYG does not have an established orthography, some conventions have been developed in CYG written texts for sounds that exist in the dialect but not in the Standard, which also maintain the grapheme-to-phoneme correspondence (Arvaniti, 1999; Koutsoudas & Koutsoudas, 1962; Simaki et al., 2015). According to Koda (1989), there is a strong relationship between orthography and cognition, and the strategies of phonological coding used in the processing of the L1 are transferred to the processing of the L2, meaning that CYG learners of English are likely to simply transfer their L1 strategies to the L2. Therefore, being used to the association of grapheme to phoneme, CYG learners may try to associate an L1 phoneme with graphemes that have the same form in both CYG and English, or to find an association between phonemes and graphemes, despite knowing that this does not apply to English (Koutsoudas & Koutsoudas, 1962).

Furthermore, it should also be noted that L2 vowel perception can be strongly affected by consonantal context, although this may not affect all vowels equally. Bohn and Steinlen (2003), for example, examined the identification of the 11 Standard British English monophthongs by Danish listeners in 3 different contexts (/hVt/, /dVt/ and /gVt/) and found that the perceptual assimilation of /ɪ ɛ ʊ ʌ/ was strongly affected by consonantal context, while /ɔː ɜː/ were not much affected by context and /iː uː æ ɒ ɑː/ were very consistently identified across contexts. At the same time, such effects are expected to occur at the initial stages of acquisition, since more experienced learners were found to disregard contextual variation and perceive L2 vowels more consistently (e.g. Balas, 2018; Levy & Strange, 2008). Since the present research is concerned with CYG learners in FLA contexts, with limited exposure to native-like input, it is predicted that their perception and production of the target L2 vowels will be affected by consonantal context. However, in the absence of any previous studies investigating this effect in learners of the same or similar L1 backgrounds and L2 experience, it is difficult to make more concrete predictions as to which vowels might be affected to a larger extent.

Therefore, the current research aimed to address the following research questions, in an attempt to provide an in-depth examination of the realisation of vowels in one of the most widely-spoken, yet largely unstudied, varieties of SMG, as well as the acquisition of L2 English vowels by native speakers of CYG:

1. How do adult CYG learners perceive and produce L2 English vowels compared to native English (NE) speakers?
2. Is learners' production of L2 English vowels affected by orthographic cues?
3. Is learners' perceptual or production performance affected by consonantal context?

Given that the vowel inventory of the Cypriot-Greek (CYG) dialect is considerably smaller than the English vowel inventory and does not use spectral or durational cues to signal phonological contrast (Coutsougera, 2007; Themistocleous & Logotheti, 2016), these learners of English may face difficulties in successfully using these cues to detect the differences between contrasting vowels which overlap a single L1 category (Best, 1995; Best & Tyler, 2007; Flege, 1995). As a result, learners are likely to produce contrasting vowels without the required durational or spectral differences, merging them in a single, L1-based category. At the same time, L2 vowel perception and production may be further hindered by the orthographic differences between the two languages (Stoehr & Martin, 2022) as well as consonantal context (Bohn & Steinlen, 2003).

3. METHODS

3.1. Participants

Fourteen native CYG (7 male, 7 female) and 8 NE speakers (3 male, 5 female) were recruited for the purposes of this study. All participants were volunteers and the procedure adhered to all ethical standards. None of the participants reported having a speech or hearing impairment. Participants were 18–28-year-old students in various disciplines and various years of study at an English-speaking university in Cyprus or the UK, respectively.

3.2. Stimuli

The stimuli used in the study included the 11 SSBE target vowels (/ɪ iː e ɜː ɑː æ ʌ ɒ ɔː ʊ uː/) in the CVC contexts /bVt/ and /gVt/ (or /gVd/), in order to enable the examination of context effects: bit/git, beat/*gheat, bet/get, *burt/*gert, *bart/*gart, bat/gat, but/gut, bot/got, bought/*gort, butch/good, boot/*gould, respectively (adapted from Bohn & Bundgaard-Nielsen, 2008; Lengeris, 2009; Mayr & Escudero, 2010). CYG speakers also produced the five CYG vowels in a /'bVtv/ context, where only the stressed vowel was analyzed: */'bita/, */'beta/, */'bata/, /'bota/, */'buta/ (adapted from Papachristou, 2011). A disyllabic structure was preferred in order to ensure that the target vowels appeared in phonotactically permitted sequences. Both real and non-words (indicated by asterisks) were included in the stimuli in order to ensure that the vowels appeared in matching environments in both languages (Di Paolo et al., 2011; Kerswill & Watson, 2014), minimise potential word-frequency effects (Huensch & Tremblay, 2015) and direct learners' attention on phonetic form and not meaning (Thomson, 2011). Since the emphasis was placed on ensuring that the vowels appeared in matching environments, real and non-words were not matched, similarly to other studies (e.g. Bohn & Bundgaard-Nielsen, 2008; Lengeris, 2009; Mayr & Escudero, 2010).

3.3. Procedure

The procedure is outlined in Fig. 1 below. The following sections provide further details about each completed task.

3.4. Perceptual tasks

CYG participants completed an 11-alternative forced-choice identification (FCID) task without feedback, in which they heard a total of 148 stimuli through headphones. Participants could listen to each stimulus up to 9 times before selecting their response by clicking on the label containing the word they heard on a computer screen. The task was administered through the TP (Teste/Treino de Percepção – Perception Testing/Training) software (Rauber et al., 2012) and used stimuli isolated from NE speakers' productions in Recording 1 (see section 3.5). Each target word with its corresponding label was introduced before the test. Similarly to Bohn and Bundgaard-Nielsen (2008) the label "bUt" was used for the stimulus "butch" to differentiate it from "but". In order to avoid confusion, participants were instructed in

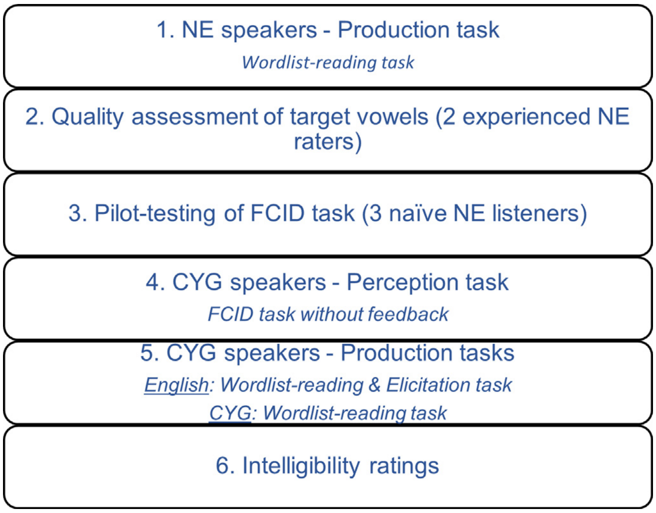

Fig. 1. Data collection procedure.

advance about this label; during the presentation of labels with their corresponding target words, it was stressed that the label "but" corresponded to the real word as they know it, whereas the label "bUt" corresponded to the word "butcher" and rhymed with "good". All occurrences of this target word were manipulated to /bʊt/, so that context would not facilitate the identification of the target word. An "Oops" button was enabled, allowing participants to choose a different label if they accidentally clicked on the wrong button in their immediately previous response. Twenty randomly selected items were added at the beginning for familiarization. The average time to complete the perceptual task was 19 min (range: 12–29 min).

Before administering the perceptual task, it was also completed by 3 NE naïve listeners to ensure its validity and reliability. Listeners achieved an average accuracy rate of 87.86 % (range: 85.07 %-89.55 %). Stimuli that were wrongly identified by two or more raters were removed from the perceptual task and all other analyses. In agreement with Thomson (2011), the remaining stimuli that were misidentified by one listener only were not considered problematic, as they represented a very small proportion of the stimuli, they reflected the variation found in naturalistic speech, and their misidentification may also have been due to individual listener characteristics.

3.5. Production tasks

Production data were collected in quiet university rooms or labs in the UK and Cyprus using a Zoom H1 audio recorder (sampling rate 44.1 kHz). NE speakers produced the English stimuli, fillers and four practice items in wordlist-reading tasks with two carrier phrases: "He said. . .and left" (Recording 1) and "The next word is. . ." (Recording 2), which were visually presented in random order using PowerPoint. The target words in Recording 1 were used for analyses and Recording 2 was used in the elicitation task described below. Two NE raters trained in phonetics assessed the quality of the target vowels in terms of clarity of the recordings, voice quality and target-likeness, resulting in the exclusion of 20 stimuli from the original 352 assessed.

CYG participants produced all stimuli in a wordlist-reading and in an elicitation task to enable an examination of whether CYG learners' productions were affected by English spelling. In the wordlist-reading task, the target words along with fillers were embedded in the carrier phrase "He said. . .and left", which were visually presented to participants in random order using PowerPoint. In the elicitation task, CYG participants listened to the target words and fillers presented in the carrier phrase "The next word is. . ." via headphones, and repeated them in a new carrier phrase "Now I say. . .for you" (similarly to Thomson & Derwing, 2016). Finally, CYG speakers produced the CYG stimuli twice in a wordlist-reading task at a normal speaking rate. These productions enabled the acoustic analysis of CYG vowels in order to measure the cross-linguistic similarity or dissimilarity of vowels in the two languages, which could possibly explain CYG learners' perceptual and production patterns. The target words and fillers were embedded in the carrier phrase /ˈipendu. . .ˈdʒefie/, ("He said to him. . .and left") presented in random order in PowerPoint. Each task included 4 practice items at the beginning and was recorded twice, although the best production was chosen for analysis, as determined by the quality of the recording, the quality of the target vowel (most target-like), voice quality, hesitation in producing the target word or mispronunciation of the target word.

3.6. Analysis

Perceptual performance was measured through %-correct identification scores and statistical analyses were conducted using mixed-effects binomial logistic regression with dummy coding in R (Version 4.2.1), with the significance level set at 0.05. Production data were analyzed acoustically using Praat (Version 6.1.37) and through intelligibility ratings by NE listeners. The target vowels were segmented manually, and duration and formant values were measured using simultaneous inspections of the waveform and spectrogram. Vowel onset and offset were identified based on the onset and offset of periodic energy in F2 and higher formants in the spectrogram as shown in Fig. 2 (Nishi et al., 2008; Themistocleous, 2017b). Formant settings were adjusted for individual speakers and F1 and F2 values were measured at mid-point (Hutchinson & Dmitrieva, 2022). Outliers in duration and formant values were identified through data visualization with a box plot in MS Excel and were removed before proceeding with any statistical analyses. Effects of physiological differences between male and female participants were removed using the Lobanov (1971) method of normalization.

Vowel overlap was measured by calculating the Pillai score (Hay et al., 2006) through MANOVA tests in R. Higher Pillai scores indicate a smaller amount of vowel overlap whereas lower Pillai scores indicate a higher amount of overlap between two vowels (Hay et al., 2006). Statistical analyses on duration were conducted in R, using the lme4 package (Version 1.1–30) (Bates et al., 2015), with the significance level set at 0.05. A 10 % representative sample of the recordings was submitted to intra-rater reliability assessment (Mildner, 2013). The two measurements did not differ significantly in either duration, as determined by a paired-samples *t*-test (p = 0.143), or spectral characteristics, as shown by the high degree of overlap between them (highest Pillai score obtained: 0.004).

Intelligibility ratings were included to complement acoustic analyses, assess how L2 productions are evaluated by NE listeners and determine whether any adjustments made by the learners were perceptually salient (Hutchinson & Dmitrieva, 2022). Naïve listeners were preferred in order to examine how the speech of L2 learners is evaluated in naturalistic conditions that resemble real-life situations. Five native SSBE speakers aged between 18–28 years old served as the raters. They reported that they did not speak any other languages or have a speech or hearing impairment. Before proceeding to the task, they completed a familiarization task, which included the target words produced by 2 NE speakers (1 male, 1 female) and this task had to be completed without errors.

The raters provided intelligibility measurements of the target vowels in their context through an 11-alternative FCID task without feedback. Following Flege and Wayland (2019), the stimuli were edited so that consonant production would not influence listeners' ratings of the target vowels: any pre-voicing in word-initial voiced stops and any instances of post-vocalic /r/ or /l/ were digitally removed, and any instances of word-final /t/ or /d/ affected by the preceding approx-

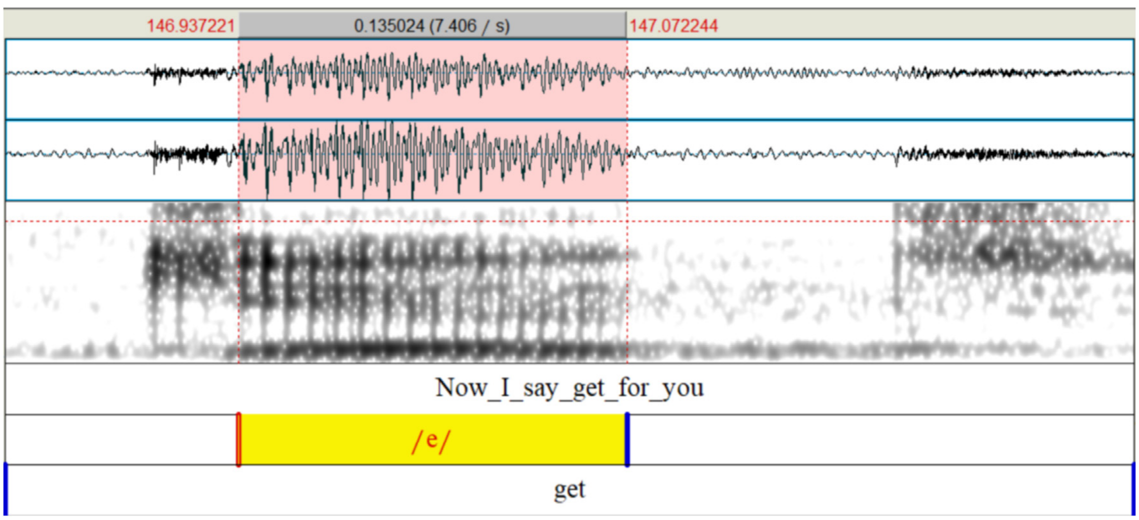

Fig. 2. Example of segmentation.

imant were replaced by portions of the same participant's productions of word-final /t/ or /d/. All stimuli were checked for naturalness before administering the task.

Each listener rated the whole set of target words produced by learners which were subdivided into 10 smaller tasks containing randomly selected stimuli across participants and tasks, blocked by context and presented in random order. Raters completed no more than one task per day, in order to avoid the effects of raters' fatigue (Suzukida & Saito, 2019). Inter-rater agreement was established through Cronbach's alpha (α = 0.82). Raters' identification scores were analyzed using mixed-effects binomial logistic regression in R.

For ease of reference and to make the results of this study comparable to others that have chosen other symbols to represent the target vowels, the remainder of this paper will refer to the target vowels using Wells' (1982) keywords, as follows: KIT for /ɪ/, FLEECE for /iː/, DRESS for /e/, NURSE for /ɜː/, BATH for /ɑː/, TRAP for /æ/, STRUT for /ʌ/, LOT for /ɒ/, NORTH for /ɔː/, FOOT for /ʊ/ and GOOSE for /uː/.

4. RESULTS

4.1. Perception

Tables 1 and 2 present the confusion matrices showing participants' percentage of responses to each target vowel and the most commonly confused vowels. Contextual effects were evident in the confusion patterns of target vowels as well, although most vowel pairs confirmed to a large extent the prediction that contrastive vowels would be confused with each other, with the exception of the DRESS-NURSE pair. More specifically, in the /bVt/ context, DRESS was challenging and mostly confused with STRUT instead of the expected NURSE, whereas NURSE was relatively better perceived and also occasionally confused with STRUT. Conversely, in the /gVt/ context, DRESS was well-perceived with occasional confusion with NURSE, as expected, whereas NURSE was poorly perceived and confused with BATH and STRUT instead of the expected DRESS, eliciting a variety of responses.

CYG learners' overall perceptual identification score in the two contexts was similar (/bVt/: 39.96 %, /gVt/: 37.71 %), and binomial logistic regression with *Context* as the fixed and *Participant* and *Vowel* as the random intercepts (*glmer (Identification.Score ~ Context+(1|Participant)+(1|Vowel), data = data.frame, family="binomial")*) did not detect a significant difference between them. Differences in the mean %-correct identification scores of each vowel as an effect of *Context* (Fig. 3) were assessed using binomial logistic regressions with *Participant* as the random intercept (*glmer(Identification.Score ~ Context+(1|Participant), data = data.frame, family="binomial")*) which showed significant differences in most vowels, suggesting strong contextual effects in the perception of L2 vowels: KIT (est. = -2.111, $p < 0.001$), DRESS (est. = 1.606, $p < 0.001$), NURSE (est. = -1.603, $p < 0.001$), BATH (est. = 1.592, $p = 0.001$), STRUT (est. = -1.447, $p < 0.001$), LOT (est. = 2.155, $p < 0.001$), NORTH (est. = -1.553, $p < 0.001$), FOOT (est. = 2.328, $p < 0.001$). These results demonstrate that the /bVt/ context facilitated identification for KIT, NURSE, STRUT and NORTH, while the /gVt/ context facilitated the identification of DRESS, BATH, LOT and FOOT.

Contextual effects were also observed in the identification of L2 contrastive vowels in almost all vowel pairs, since different pairs presented differing degrees of difficulty for learners depending on the context of the target vowel, as

Table 1
Confusion matrix of participants' percentage of responses to each target vowel in the /bVt/ context.

	Response										
Target vowel	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**49.1**	**<u>20.5</u>**	14.3		1.8	3.6	1.8	1.8		2.7	3.6
FLEECE	**<u>53.2</u>**	**36.9**	4.5								2.7
DRESS	1.8		**28.6**	<u>3.6</u>	5.4	25	**28.6**		1.8	5.4	
NURSE		2.4	<u>7.1</u>	**40.5**	13.1	9.5	**16.7**	2.4	1.2	4.8	2.4
BATH				14.3	9.5	**<u>21.4</u>**	**<u>33.3</u>**	7.1	7.1	6	1.2
TRAP				2.7	<u>6.3</u>	**41.4**	**<u>41.4</u>**			5.4	
STRUT	1			5.1	<u>11.2</u>	**<u>25.5</u>**	**45.9**	3.1	3.1	5.1	
LOT	1			2	9.2	3.1	7.1	**25.5**	**<u>45.9</u>**	4.1	2
NORTH				6.3				<u>12.5</u>	**50.9**	5.4	**23.2**
FOOT	3.6		14.3	3.6	7.1	3.6	10.7	7.1	7.1	**17.9**	**<u>25</u>**
GOOSE	6.3	9.8				1.8			5.4	**<u>11.6</u>**	**62.5**

Note. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%).

Table 2
Confusion matrix of participants' percentage of responses to each target vowel in the /gVt/ context.

	Response										
Target vowel	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**12.5**	8	**60.7**	9.8		1.8	1.8	1.8		2.7	
FLEECE	**53.6**	**30.4**	7.1	4.5						2.7	
DRESS	3.1	1	**63.3**	**13.3**	5.1	4.1	5.1	4.1		1	
NURSE		2.7	8	**16.1**	**16.1**	14.3	**17**	12.5	1.8	5.4	6.3
BATH	1.4	1.4		4.3	**30**	10	14.3	**27.1**	8.6		2.9
TRAP	1	1	1	1	**21.4**	**44.9**	17.3	8.2	2		2
STRUT	1.2	2.4		7.1	19	**29.8**	16.7	**20.2**	1.2	1.2	1.2
LOT				1.8	1.8	2.7	5.4	**66.1**	**16.1**	2.7	1.8
NORTH					1.8		1.8	**29.7**	19.8	**27.9**	16.2
FOOT			7.1	4.5		1.8	6.3	4.5		**67.9**	4.5
GOOSE			8.9						6.7	**33.3**	**51.1**

Note. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%).

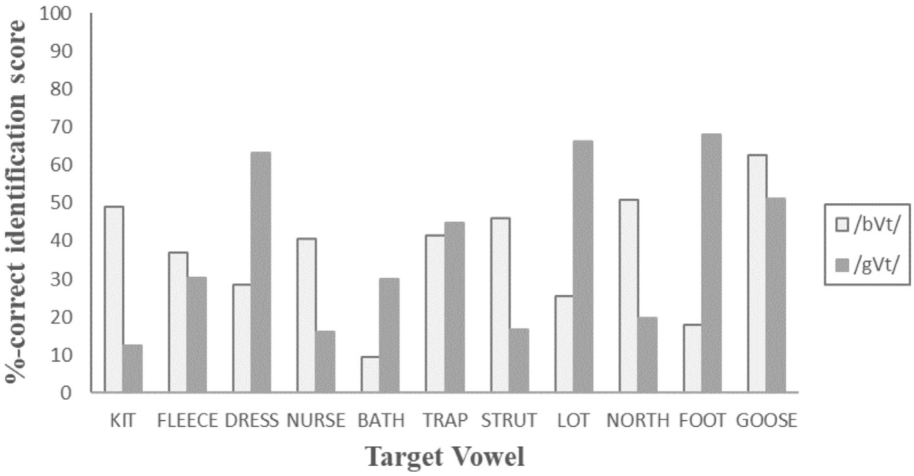

Fig. 3. Mean perceptual identification scores of each target vowel in each context.

shown by a series of binomial logistic regressions with *Vowel* as the fixed and *Participant* as the random effects conducted for each vowel contrast (*glmer(Identification.Score ~ Vowel+(1|Participant), data = data.frame, family=*"*binomi al*"*)*). In the /bVt/ context, the results showed differences in the identification between the contrastive vowels BATH-TRAP (est. = 1.993, $p < 0.001$), BATH-STRUT (est. = 2.088, $p < 0.001$), LOT-NORTH (est. = 1.153, $p < 0.001$) and FOOT-GOOSE (est. = 2.057, $p < 0.001$), while in the /gVt/ context, significant differences were found between KIT-FLEECE (est. = -1.214, $p < 0.001$), DRESS-NURSE (est. = -2.447, $p < 0.001$), BATH-TRAP (est. = 0.664, $p = 0.049$), TRAP-STRUT (est. = 1.405, $p < 0.001$) and LOT-NORTH (est. = -2.232, $p < 0.001$).

4.2. Duration

Fig. 4 presents the average duration of each vowel as produced by CYG learners and NE speakers in each task. For the purposes of these analyses, the mean duration of each target vowel includes productions in both the /bVt/ and /gVt/ contexts, given the absence of any significant differences between them as determined through linear mixed-effects analyses with *Context* as the fixed and *Participant* and *Vowel* as the random effects (est. = 2.11, $p = 0.438$ in the elicitation task; est. = 1.229, $p = 0.621$ in the wordlist-reading task) (*lmer(Duration ~ Context+(1|Participant)+(1|Vowel), data = data.frame)*). As shown in Fig. 4, vowels produced in the wordlist-reading task were consistently longer than vowels in the elicitation task. This was further confirmed through a linear mixed-effects analysis with *Task* as the fixed effect

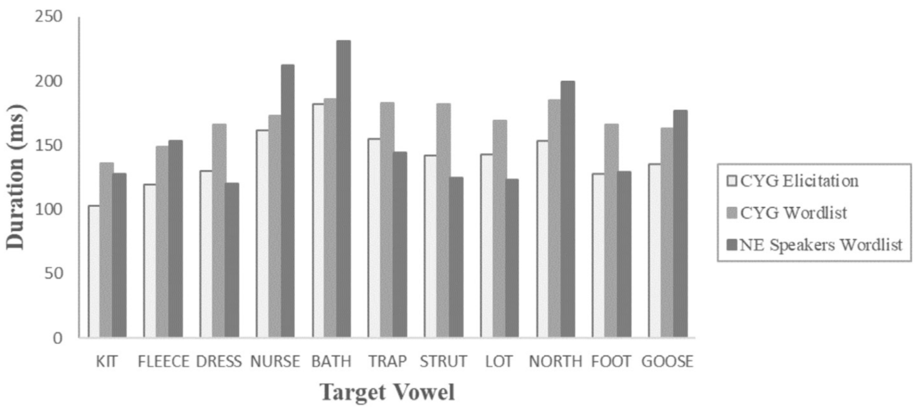

Fig. 4. Mean duration (ms) of each target vowel produced in both contexts by CYG learners and NE speakers in each task.

and *Participant*, *Context* and *Vowel* as the random effects (est. = 29.355, $p < 0.001$) (*lmer(Duration ∼ Task+(1|Participant)+(1|Context)+(1|Vowel), data = data.frame)*).

To assess whether the vowels in a vowel pair were produced with different lengths in each of the tasks, the effect of *Vowel* on duration was examined through linear-mixed effect analyses for each vowel pair in each task, with *Participant* as the random effect (*lmer(Duration ∼ Vowel+(1|Participant), data = data.frame)*). In the elicitation task, vowel length differences were observed in KIT-FLEECE (est. = -15.74, $p = 0.016$), DRESS-NURSE (est. = 29.27, $p < 0.001$), BATH-TRAP (est. = –23.004, $p = 0.002$), BATH-STRUT (est. = -39.195, $p < 0.001$) and TRAP-STRUT (est. = 15.871, $p = 0.005$). In the wordlist-reading task, vowel length differences were found only in KIT-FLEECE (est. = -13.34, $p = 0.007$) and LOT-NORTH (est. = 16.16, $p < 0.001$).

Differences between NE speakers and CYG learners in the length of the target vowels were examined through linear mixed-effects analyses for each target vowel, with *Group* as the fixed and *Participant* as the random effect (*lmer(Duration ∼ Group+(1|Participant), data = data.frame)*). In the wordlist-reading task, the analysis revealed significant differences between the two groups in most vowels: DRESS (est. = -47.5, $p = 0.001$), NURSE (est. = 37.518, $p = 0.002$), BATH (est. = 45.523, $p < 0.001$), TRAP (est. = -40.035, $p = 0.002$), STRUT (est. = -59.728, $p < 0.001$), LOT (est. = -46.373, $p < 0.001$) and FOOT (est. = -41.07, $p = 0.014$). Even though NE speakers only produced the vowels in a wordlist-reading task, the same analysis was conducted comparing their productions with CYG learners' productions in the elicitation task. This analysis showed significant differences in KIT (est. = 24.116, $p = 0.025$), FLEECE (est. = 32.762, $p = 0.016$), NURSE (est. = 47.413, $p = 0.002$), BATH (est. = 49.625, $p < 0.001$), NORTH (est. = 44.255, $p < 0.001$) and GOOSE (est. = 42.678, $p = 0.002$), showing that the shorter durations of vowels as produced by CYG learners in the elicitation task approximate the durations of all short vowels as produced by the NE speakers, with the exception of KIT. However, the differences in long vowels may also be attributed to a task effect that was not further examined.

Finally, the duration of CYG vowels was as follows: /i/=114 ms, /e/=126 ms, /a/=137 ms, /o/=132 ms and /u/=117 ms. An one-way ANOVA (*aov(Duration ∼ Vowel, data = data.frame)*) demonstrated no significant effect of *Vowel* on duration, suggesting that the CYG vowels as produced by these speakers were not differentiated by duration. CYG vowels were shorter than L2 long vowels and approximated the durations of the corresponding L2 short vowels, as confirmed through linear mixed-effects model analysis where each CYG vowel was compared with the corresponding L2 vowels produced by NE speakers in separate models (FLEECE-i, KIT-i, DRESS-e, NURSE-e, BATH-a, TRAP-a, STRUT-a, LOT-o, NORTH-o, FOOT-u, GOOSE-u) with *Vowel* as the fixed and *Participant* as the random effects (*lmer(Duration ∼ Vowel+(1|Participant), data = data.frame)*). A significant effect of *Vowel* was observed in the comparisons between each CYG vowel and the corresponding long L2 vowel (FLEECE-i: est. = -39.566; NURSE-e: est. = 85.087; BATH-a: est. = 94.421; NORTH-o: est. = -65.991; GOOSE-u: est. = -60.673; $p < 0.001$ in all comparisons), whereas no effect was found in the comparisons of CYG vowels with the corresponding short vowels KIT, DRESS, TRAP, STRUT, LOT and FOOT.

4.3. Spectral characteristics

Table 3 shows the mean formant values of each vowel in each context as produced by CYG learners in both tasks, and the Pillai scores obtained for each comparison between them (*manova(cbind(F1,F2) ∼ Context, data = data.*

Table 3
Mean, Lobanov-normalised F1 and F2 values (Hz) of each vowel in each context (both tasks), overlap between them (Pillai score) and significance of result.

Vowel	Context				Pillai Score	*p*
	/bVt/		*/gVt/*			
	F1	*F2*	*F1*	*F2*		
KIT	306	1769	305	1788	0.0044	
FLEECE	301	1746	289	1822	0.1588	0.017
DRESS	422	1545	402	1595	0.2342	<0.001
NURSE	419	1391	411	1488	0.2072	0.019
BATH	503	1175	520	1227	0.0569	
TRAP	551	1261	544	1315	0.116	0.046
STRUT	533	1234	514	1289	0.1657	0.013
LOT	416	1039	421	1087	0.1476	0.018
NORTH	400	1027	403	1047	0.0223	
FOOT	326	1077	321	1144	0.1165	
GOOSE	321	1077	327	1017	0.0961	

frame)). Fig. 5 shows the plotted vowels for each of the contexts, demonstrating the amount of overlap between them. As can be seen, there is a high degree of overlap in the vowels produced in each context, and therefore, analyses were conducted with combined /bVt/ and /gVt/ mean values.

Table 4 shows the mean F1 and F2 values of each vowel in each task as produced by CYG learners in both contexts and the Pillai scores obtained for each comparison between them (*manova(cbind(F1,F2) ~ Task, data = data.frame)*). Fig. 6 shows the plotted vowels for each of the tasks, demonstrating the high degree of overlap in the vowels produced in each task.

The examination of the vowels in a vowel pair in each task (*manova(cbind(F1, F2) ~ Vowel, data = data.frame)*) showed considerable overlap between contrastive vowels (Table 5), with the highest Pillai score found in DRESS-NURSE in the elicitation task. Given some overlap observed between BATH and LOT, the Pillai score was also calculated for this pair.

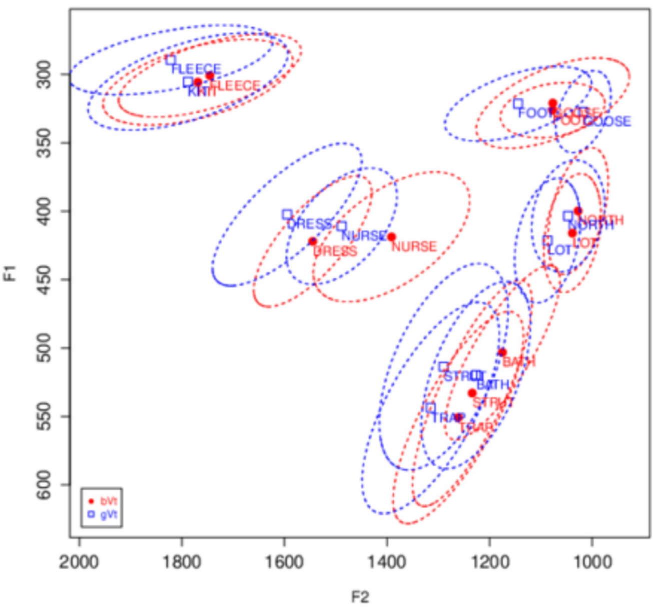

Fig. 5. F1xF2 plot of target vowels in /bVt/ and /gVt/ contexts.

Table 4
Mean, Lobanov-normalised F1 and F2 values (Hz) of each vowel in each task (both contexts), overlap between them (Pillai score) and significance of result.

Vowel	Task				Pillai Score	*p*
	Elicitation		*Wordlist*			
	F1	*F2*	*F1*	*F2*		
KIT	309	1756	302	1801	0.059	
FLEECE	290	1760	300	1811	0.038	
DRESS	405	1568	419	1571	0.035	
NURSE	409	1405	420	1466	0.0485	
BATH	494	1189	524	1211	0.055	
TRAP	541	1314	554	1260	0.1534	0.016
STRUT	505	1270	541	1254	0.123	0.043
LOT	415	1076	422	1049	0.0687	
NORTH	396	1052	408	1022	0.103	
FOOT	327	1112	319	1121	0.0358	
GOOSE	330	1108	317	1007	0.1668	0.022

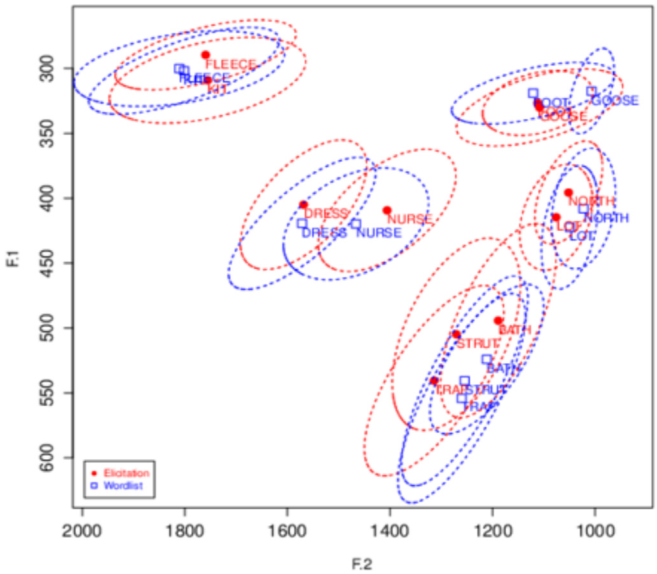

Fig. 6. F1xF2 plot of the overlap of target vowels produced by CYG learners in the wordlist-reading and elicitation tasks.

Table 5
Overlap between vowels of contrastive pairs as produced by CYG learners.

Vowel Pair	Elicitation		Wordlist	
	Pillai Score	*p*	Pillai Score	*p*
KIT-FLEECE	0.1276	0.033	0.0031	
DRESS-NURSE	0.3994	<0.001	0.1827	0.011
BATH-TRAP	0.1973	0.011	0.0499	
BATH-STRUT	0.1225		0.0349	
TRAP-STRUT	0.0581		0.012	
LOT-NORTH	0.0621		0.0583	
FOOT-GOOSE	0.0063		0.238	
BATH-LOT	0.3677	<0.001	0.5389	<0.001

Table 6
Mean Lobanov-normalised F1 and F2 (in Hz) of English vowels produced by NE speakers.

Vowel	F1	F2
KIT	376	1717
FLEECE	297	1973
DRESS	509	1603
NURSE	463	1407
BATH	475	1166
TRAP	548	1409
STRUT	513	1302
LOT	434	1132
NORTH	349	1001
FOOT	364	1421
GOOSE	313	1427

Table 6 lists the mean F1 and F2 values of each vowel produced by NE speakers in the wordlist-reading task in both contexts, which are plotted in Fig. 7 as well.

The overlap between vowels in a contrastive pair for NE speakers is shown in Table 7 (*manova(cbind(F1,F2) ~ Vowel, data = data.frame)*). As opposed to CYG learners, the vowel productions of NE speakers were more widely spread in the vowel space, and there was less overlap between the members of a contrastive pair.

The overlap between the productions of NE speakers and CYG learners in each task is shown in Table 8 (*manova (cbind(F1,F2) ~ Group, data = data.frame)*). Fig. 8 shows the plotted vowels for each task compared to NE productions. As evident by these results, learners produced the target vowels differently than NE speakers. Notably, KIT was much

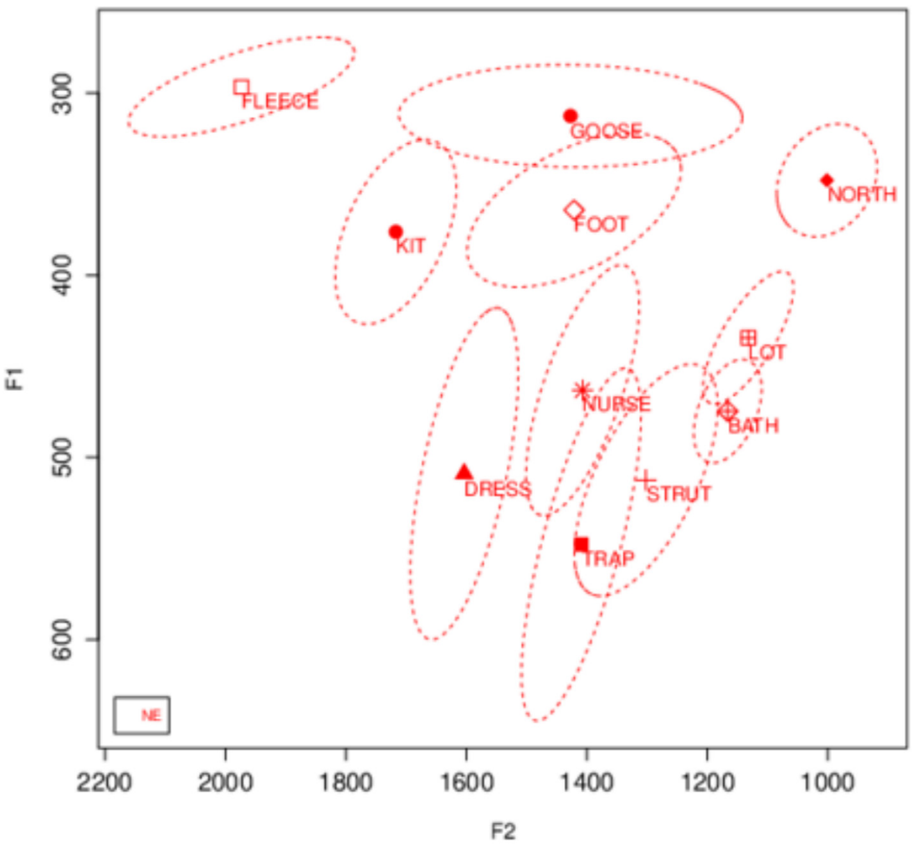

Fig. 7. F1xF2 plot of vowels produced by NE speakers.

Table 7
Overlap between vowels of a vowel pair in NE speakers' productions.

Vowel Pair	Pillai score	*p*
KIT-FLEECE	0.7765	<0.001
DRESS-NURSE	0.5933	<0.001
BATH-TRAP	0.7304	<0.001
BATH-STRUT	0.3468	0.002
TRAP-STRUT	0.2185	0.013
LOT-NORTH	0.6387	<0.001
FOOT-GOOSE	0.3576	<0.01

Table 8
Overlap between NE speakers and CYG learners' vowel productions in each task.

Vowel	Elicitation		Wordlist	
	Pillai Score	*p*	Pillai Score	*p*
KIT	0.4763	<0.001	0.6248	<0.001
FLEECE	0.3169	<0.001	0.2147	0.006
DRESS	0.4015	<0.001	0.352	<0.001
NURSE	0.2439	0.01	0.2878	0.002
BATH	0.0295		0.2083	0.012
TRAP	0.2061	0.007	0.5182	<0.001
STRUT	0.0165		0.2176	0.007
LOT	0.1344		0.3167	<0.001
NORTH	0.3027	<0.001	0.3518	<0.001
FOOT	0.5518	<0.001	0.4719	<0.001
GOOSE	0.4315	<0.001	0.5826	<0.001

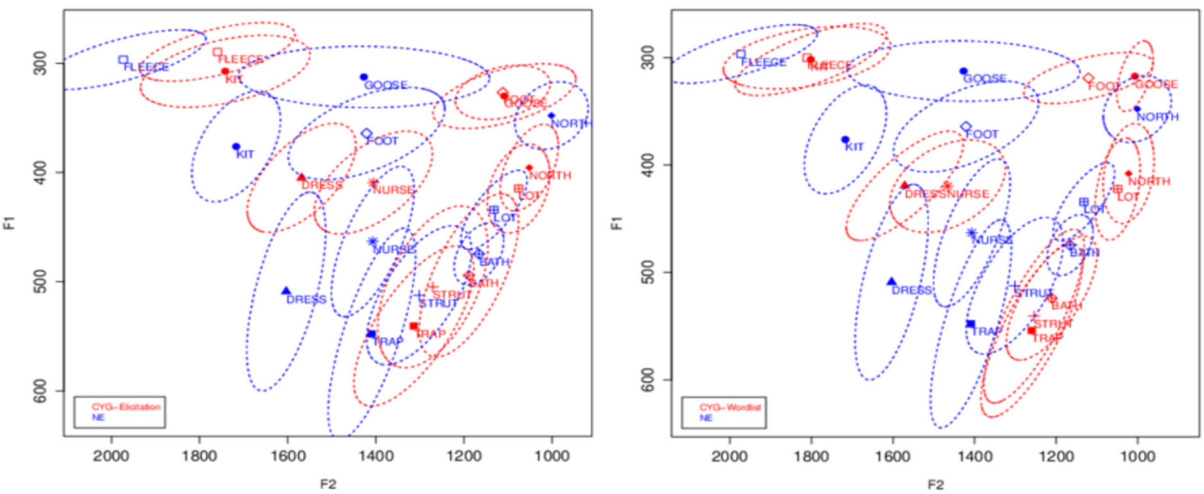

Fig. 8. F1xF2 plots of NE speakers and CYG learners' (both groups) productions in the elicitation (left) and wordlist-reading (right) tasks.

lower in the productions of NE speakers, whereas FOOT and GOOSE were much fronter, as is common in contemporary SSBE. TRAP also had a low degree of overlap with NE speakers' productions in the wordlist-reading task, as opposed to the elicitation task where some overlap was observed. On the other hand, BATH and STRUT had a high degree of overlap between learners and NE speakers, particularly in the elicitation task. However, in the case of BATH, the high degree of overlap can be at least partly attributed to the large variation found in learners' productions, which were more widely spread, taking up more space than NE speakers' productions, who produced the two vowels with less overlap between them.

Table 9 shows the mean Lobanov-normalised formant values for CYG vowels. The Pillai scores showing the overlap between the L1 and corresponding L2 vowels produced by the learners in each task are shown in Table 10 (*manova (cbind(F1,F2) ~ Vowel, data = data.frame)*), and the respective F1xF2 plots are shown in Fig. 9.

These analyses demonstrate that CYG learners produced the L2 vowels clustered around their L1 vowels. The most deviation from L1 articulatory routines was observed in NURSE and BATH in the elicitation task, and in NORTH, FOOT and GOOSE in the wordlist-reading task, although these still had a high degree of overlap with the respective contrastive L2 vowels and the L1 vowels.

Finally, the CYG and SSBE vowels as produced by the respective native speakers of each variety were compared in an attempt to explain the performance of learners in L2 vowel production and their differences with NE speakers. The Pillai scores showing the overlap between CYG and SSBE vowels are shown in Table 11 (*manova(cbind(F1,F2) ~ Vowel, data = data.frame)*), and the vowels are plotted in Fig. 10.

As shown in Fig. 10, CYG /i/ is found between English FLEECE and KIT, with almost equal distance from each. As opposed to this, CYG /e/ and CYG /o/ are closer to the corresponding short L2 vowels DRESS and LOT, respectively, than the long L2 vowels NURSE and NORTH, respectively. Furthermore, CYG /u/ is almost equally distant from the L2 vowels GOOSE and FOOT, which are considerably fronter, and is closer to the back vowel NORTH. Finally, CYG /a/ has more overlap with L2 STRUT than BATH or TRAP.

4.4. Intelligibility ratings

Fig. 11 shows the mean intelligibility scores of the target vowels as rated by NE raters in each task for both contexts. Overall, the most intelligible vowels produced by CYG learners according to the ratings were FLEECE, and the members of the DRESS-NURSE contrast, while the most problematic vowels were STRUT and KIT. Further analyses to examine the effects of *Task* and *Vowel* on the intelligibility of target vowels were conducted using binomial logistic regression to assess how the productions of CYG learners were perceived by NE raters, in order to complement the results of acoustic analyses. The effect of *Task* (*glmer(Result ~ Task+(1|Participant)+(1|Rater), data = data.frame, family="binomial")*) reached significance in three vowels: KIT (est. = -0.992, p = 0.015), which was found to be more intelligible in the elicitation task, and BATH (est. = 0.783, p = 0.019) and NORTH (est. = 0.908, p = 0.002), which were found to be more intelligible in the wordlist-reading task. The effect of *Vowel* (*glmer(Result ~ Vowel+(1|Participant)+(1|Rater), data = data.frame, family="binomial")*) reached significance in all vowel pairs examined in both tasks, with the exception of the FOOT-GOOSE contrast, as shown in Table 12, demonstrating that one of the vowels in each of these pairs was significantly more intelligible than the other.

Tables 13 and 14 present the confusion matrices with NE raters' percentage of responses to each target vowel and the most commonly confused vowels as produced by all CYG learners in the elicitation and wordlist-reading tasks, respectively. It is evident from these that the predicted confusion patterns are for the most part confirmed by the responses of the raters. One noteworthy exception is the case of TRAP, which was expected to be mostly confused with BATH and/or STRUT, but was instead mostly perceived as LOT by NE raters in both tasks. The same can be observed for BATH, although in the wordlist-reading task, this vowel was better perceived as the correct vowel, followed by the expected TRAP and then LOT.

5. DISCUSSION

5.1. Perception and production of contrastive vowels

Overall, both vowels in the pairs KIT-FLEECE and LOT-NORTH were challenging in perception, while the intelligibility of one of the vowels in each pair (i.e. FLEECE and LOT) as rated by NE listeners was significantly higher than

Table 9
Mean Lobanov-normalised formant values for CYG vowels.

Vowel	F1	F2
i	309.6	1747.7
e	426.2	1542.7
a	565.1	1281.6
o	419.1	1083.7
u	333.9	1054.2

Table 10
Pillai score of the productions of L1 vowels and the corresponding L2 vowel by CYG learners in each task and significance of result.

Vowels	Elicitation		Wordlist	
	Pillai Score	*p*	Pillai Score	*p*
KIT_i	0.001		0.1014	
FLEECE_i	0.2046	0.016	0.0915	
DRESS_e	0.1309		0.0578	
NURSE_e	0.2613	0.014	0.0859	
BATH_a	0.2574	0.013	0.1346	
TRAP_a	0.1139		0.0099	
STRUT_a	0.1695	0.035	0.0251	
LOT_o	0.0054		0.1073	
NORTH_o	0.1025		0.2104	0.014
FOOT_u	0.1388		0.2757	0.008
GOOSE_u	0.0793		0.2435	0.009

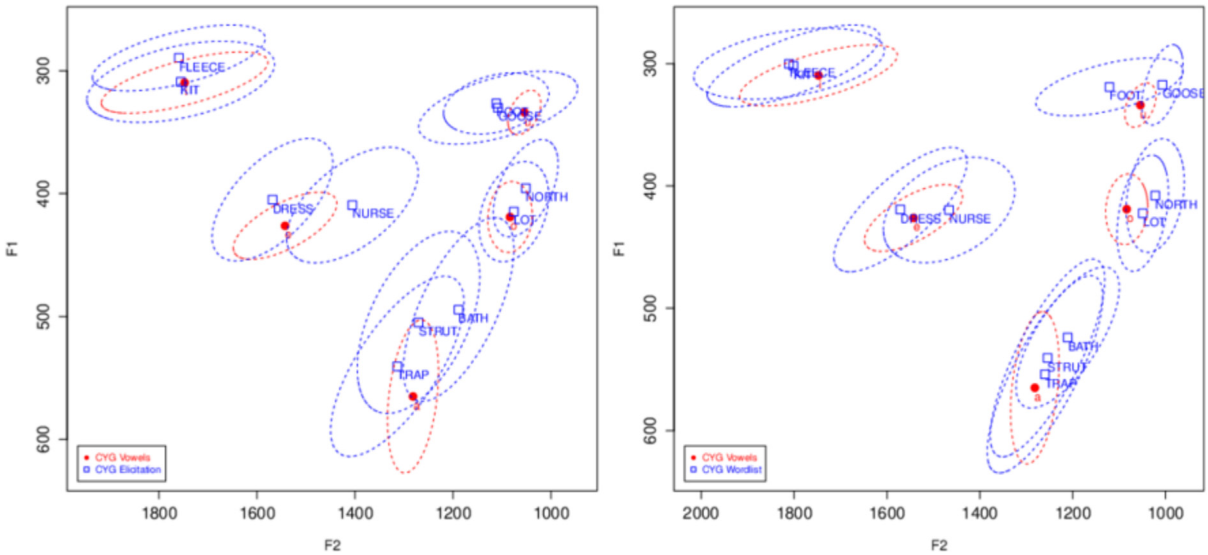

Fig. 9. F1xF2 plots of CYG vowels and L2 vowels produced in the elicitation (left) and the wordlist-reading (right) task.

Table 11
Overlap between English and corresponding CYG vowels.

Vowels	Pillai Score	*p*
KIT_i	0.4913	<0.001
FLEECE_i	0.5613	<0.001
DRESS_e	0.276	0.013
NURSE_e	0.5885	<0.001
BATH_a	0.6208	<0.001
TRAP_a	0.557	<0.001
STRUT_a	0.2494	0.016
LOT_o	0.126	
NORTH_o	0.6054	<0.001
FOOT_u	0.6502	<0.001
GOOSE_u	0.5107	<0.001
NORTH_u	0.2427	0.015

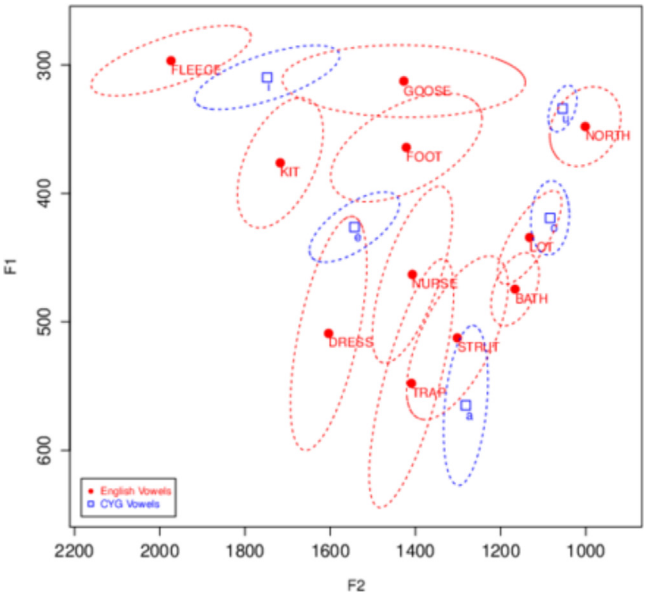

Fig. 10. F1xF2 plot of English vowels produced by NE speakers and CYG vowels.

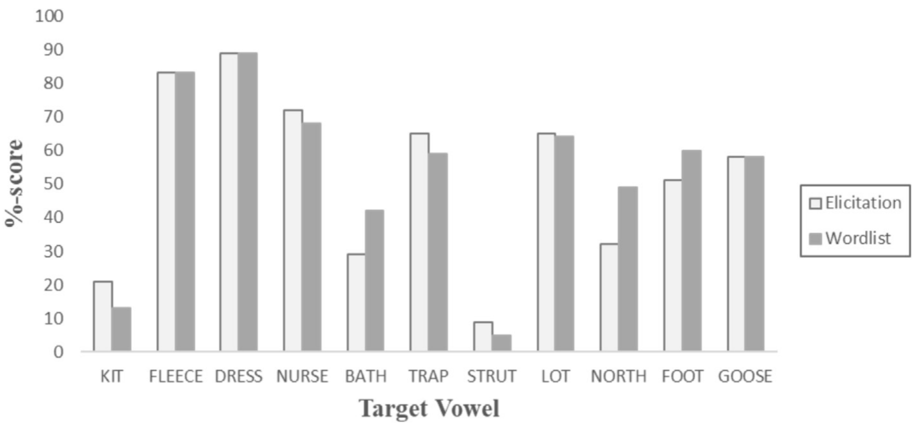

Fig. 11. Mean intelligibility scores of each target vowel in each task.

Table 12
Results of binomial logistic regression on the effect of Vowel on intelligibility score.

Vowel Pair	Elicitation		Wordlist	
	Estimate	*p*	*Estimate*	*p*
KIT-FLEECE	−3.006	<0.001	−4.101	<0.001
DRESS-NURSE	−1.203	<0.001	−1.795	<0.001
BATH-TRAP	2.065	<0.001	0.962	0.001
BATH-STRUT	−1.443	<0.001	−2.608	<0.001
TRAP-STRUT	3.093	<0.001	3.816	<0.001
LOT-NORTH	−1.467	<0.001	−0.593	0.018
FOOT-GOOSE	0.27	0.3	−0.066	0.806

Table 13
Confusion matrix of NE raters' percentage of responses to CYG learners' productions of each target vowel in the elicitation task.

	Response										
Target vowel	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**21.4**	**69.3**	8.6								
FLEECE	**12.6**	**83**	1.5								2.2
DRESS	1.4		**88.6**	4.3				1.4			
NURSE	2.4		**15.3**	**71.8**	1.2		1.2	1.2		4.7	2.4
BATH				5.6	**28.9**	14.4	7.8	**38.9**	4.4		
TRAP			2.2	3.7	9.6	**65.2**	5.2	**14.1**			
STRUT			3.2	5.6	**11.2**	**45.6**	8.8	24	1.6		
LOT							5.4	**65.4**	**23.1**	2.3	2.3
NORTH							3.7	**57**	**31.9**	2.2	4.4
FOOT				1.4				1.4	1.4	**51.4**	**43.6**
GOOSE				1			1	1.9	3.8	**34.3**	**58.1**

Note. *Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%).*

Table 14
Confusion matrix of NE raters' percentage of responses to CYG learners' productions of each target vowel in the wordlist-reading task.

	Response										
Target vowel	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**12.9**	**85**									1.4
FLEECE	**11.5**	**83.1**	2.3		1.5						1.5
DRESS			**88.6**	7.9						1.4	
NURSE			**28**	**68**	1	1		2			
BATH				3.1	**42.3**	**25.4**	3.1	24.6	1.5		
TRAP					13.3	**58.5**	3	**23**	1.5		
STRUT					10.8	**54.6**	5.4	**26.2**	1.5		
LOT				1.5		1.5		**63.7**	**30.4**		
NORTH							1.5	**42.3**	**49.2**	2.3	1.5
FOOT				1.9						**60**	**38.1**
GOOSE									4.8	**35.2**	**58.4**

Note. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%).

the other, since the L1 vowel used by CYG learners in production was closer to the respective L2 vowel. This is in line with previous studies (Cebrian, 2007; Lengeris, 2009) that found that the KIT-FLEECE contrast can be very challenging for learners in both production and perception. The identification of KIT as DRESS was not surprising, given the position of KIT between CYG /i/ and /e/ in the vowel space; a similar pattern was found for some Catalan listeners in Cebrian (2006) as well as some Danish listeners in Bohn and Steinlen (2003), who also perceived English KIT as similar to their L1 /e/.

As concerns LOT-NORTH, the results are consistent with predictions that learners would struggle to perceive and produce them since they belong to an SC assimilation type (Georgiou, 2019). It is worth noting that the duration differences maintained between LOT and NORTH in the wordlist-reading task may have facilitated the identification of NORTH, since productions in this task were found to be more intelligible than in the elicitation task where length differences were not maintained. In perception, learners' performance ranged from moderate to poor; despite the fact that CYG /o/ is very close to English LOT and much lower than English NORTH, learners found it challenging to differentiate between the two in perception while also confusing NORTH with the back vowels FOOT and GOOSE, which is unsurprising given the overlap of English NORTH to CYG /u/.

A similar pattern was observed in BATH-TRAP-STRUT, where TRAP was better perceived and produced by learners, despite the higher spectral overlap between the L1 vowel /a/ and the L2 STRUT. Despite this overlap, this vowel had the poorest identification of all vowels by NE raters, possibly due to the longer duration of STRUT in the productions of learners. On the other hand, TRAP, which had less overlap with NE productions, received higher intelligibility scores, indicating that it was perceived as a better exemplar of the L2 vowel, possibly due to the similarity between TRAP and /a/ in vowel height. The high degree of overlap between NE speakers and CYG learners' production of BATH is evidently due to the fact that learners' productions were more widely spread compared to NE productions. This wide variability found in the production of BATH also led to overlap with other vowels, i.e. TRAP, STRUT and LOT, which possibly affected the identification performance of NE raters in this vowel as well, as shown by the confusion matrices.

Overall, the members of the BATH-TRAP-STRUT contrast and the L1 vowel /a/ overlapped almost completely in the productions of learners, who also evidently struggled to perceive these vowels, particularly BATH. The confusion matrices show that the vowels were mostly confused with each other, although some instances of BATH and STRUT were also perceived as LOT, presumably depending on vowel backness (i.e. more fronted productions were perceived as TRAP, and backer productions were perceived as LOT). This is consistent with the assimilation patterns observed in Lengeris (2009) for SMG adult learners, who assimilated English TRAP and STRUT to the same L1 vowel (i.e. SMG /a/), but English BATH to SMG /o/. The challenges faced by learners in this contrast are understandable, given that they have previously been classified as an SC assimilation type (Georgiou, 2019), which can be even more problematic than other types of contrasts, especially for EFL learners (Best & Tyler, 2007; Tyler, 2019), and particularly since the L2 contains three vowels that can be perceived as a single L1 category.

In the DRESS-NURSE contrast, the use of the L1 /e/ for the production of DRESS was accepted by NE raters as a good instance of the target vowel, and the small modification made for NURSE seems to have been adequate to lead to high intelligibility ratings as well in both tasks. Importantly, the differentiation of NURSE from DRESS in the productions of L2 learners was directed towards NE patterns, similarly to the findings of Dimitriou (2022), where the vowels in the DRESS-NURSE contrast had the largest Euclidean distance between them compared to other contrasts, with NURSE having the second closest production to NE productions. In perception, the two vowels were not confused with each other, as expected, but with lower vowels, particularly STRUT. This was consistent with the high proficiency young learners in Georgiou (2019), who classified NURSE as either CYG /e/ or CYG /a/. As the only L2 English vowel that was not consistently assimilated to a single L1 vowel in Georgiou (2019), and based on the results of the present study as well as Dimitriou (2022), it can be concluded that CYG learners are likely to form a new category for NURSE, provided that sufficient input is received that can help them better attune their perception and production to the specific acoustic cues of this vowel.

Finally, learners' productions of the vowels FOOT and GOOSE almost completely overlapped and were produced using the L1 articulatory gestures for /u/, suggesting that they perceived both L2 vowels as sufficiently similar to their L1 category, despite their spectral differences. Inevitably, learners' productions of both vowels were very different to NE speakers' productions, possibly due to a lack of input from this variety. Intelligibility ratings showed that the two vowels were correctly identified at a moderate, above-chance level, but there was confusion between them, indicating that raters were able to perceive a high back vowel, but could not identify which one was intended. At the same time, learners were able to differentiate the two vowels in perception better than the members of other contrasts, although identification was moderate to poor. The relatively higher perceptual performance in these vowels may be attributed to the fact that they are both spectrally very different to any CYG vowels, and therefore, learners may have been able to perceive the subtle differences between them, despite being unable to produce them accurately.

5.2. General discussion

The findings reported above illustrate the difficulties L2 learners face in perceiving and producing new phonetic categories, as suggested by previous research in the field of FLA (e.g. Dimitriou, 2019; Iverson et al., 2003; Iverson and Evans, 2007, 2009; Lengeris, 2009; Sakai, 2016). The CYG vowel inventory consists of five well-separated vowels (Arvaniti, 1999; Coutsougera, 2007; Georgiou, 2019; Themistocleous, 2017a; Themistocleous, 2017b; Themistocleous & Logotheti, 2016), while the English vowel system exhibits regional variation (Hughes et al., 2013) and it is more complex, containing tense and lax vowels and using spectral quality to differentiate between the members of a contrast (Bohn & Steinlen, 2003; Deterding, 1997; Leung et al., 2016; Rato & Carlet, 2020). It is, therefore, unsurprising that CYG learners of English may struggle to use the appropriate cues to perceive and produce L2 contrasting vowels which overlap a single L1 category.

These results are in agreement with current models of speech production and perception such as the SLM (Flege, 1995) and SLM-r (Flege & Bohn, 2021), the PAM (Best, 1995) and the PAM-L2 (Best & Tyler, 2007), as well as previous findings arguing that the relationship between the L1 and L2 sound inventories is important in identifying the difficulties faced by L2 learners (e.g. Aliaga-Garcia & Mora, 2009; Cebrian, 2019; Cebrian et al., 2019; Lengeris, 2018; Lengeris & Hazan, 2007, among others). The findings also support the argument that learners will find it more difficult to avoid using an L1 instead of an L2 sound if there is a close association between them, due to the fact that replacing old habits is more difficult than learning a new set of language habits (Koutsoudas & Koutsoudas, 1962).

CYG learners perceived L2 English vowels poorly to moderately, suggesting that they cannot easily differentiate the members of an L2 contrast, and therefore, according to the SLM (Flege, 1995), new category formation is blocked due to equivalence classification. Learners' difficulties in perceiving the subtle differences between members of a contrastive pair were expected, given that most assimilation types reported in Georgiou (2019) belong to the CG or the SC type, which, according to the PAM-L2, can be challenging for L2 learners. Furthermore, such contrasts are even less likely to be acquired in the EFL context, which provides little opportunities for naturalistic exposure (Barriuso & Hayes-Harb, 2018; Fabra & Romero, 2012; Hutchinson & Dmitrieva, 2022) especially if the foreign-accented input received does not differentiate the L2 phonemes, as is likely in the case of CYG learners (Kyprianou, 2015; Tyler, 2019).

Furthermore, in line with the assumptions of the SLM (Flege, 1995) and the SLM-r (Flege & Bohn, 2021), CYG learners produced the L2 vowels using the same articulatory routines as for L1 sounds, which were rated as good instances for some vowels but not others. The use of unmodified L1 vowels in English words can be accepted by NE listeners as good exemplars of the target vowel and may go unnoticed (Flege & Wayland, 2019), as has also been observed in other studies, such as Cebrian (2007), where NE listeners perceived the Catalan vowels /i/, /ɛ/ and /ei/ as good instances of the acoustically closest English categories /i/, /ɛ/ and /eɪ/.

As concerns the perception-production link, the results of the present study do not allow for any robust conclusions to be drawn: whereas in some vowels, low perceptual performance was associated with low intelligibility scores (e.g. KIT, BATH, STRUT, NORTH) at least in one of the two production tasks, in others, low or average perceptual performance was associated with higher intelligibility scores (e.g. FLEECE, DRESS, NURSE, TRAP, LOT). Based on the assumptions of the SLM and previous studies suggesting that accurate perception precedes accurate production (e.g. Casillas, 2019; Melnik-Leroy et al., 2022; Nagle, 2018), it could be argued that learners' overall insufficient perceptual ability hindered their production performance as well, allowing them to produce intelligible vowels only when the L1 vowel was judged to be a good exemplar of the corresponding L2 vowel. However, the case of NURSE is a notable exception supporting the opposite direction of the link between the two modalities, as in Bohn and Flege (1997); while learners demonstrated poor perceptual identification of the vowel, they were able to modify their productions towards native-like patterns that were judged as highly intelligible by NE raters. Therefore, contrary to studies reporting at least a modest relationship between the two modalities (e.g. Baker & Trofimovich, 2006; Kluge et al., 2007; Levy & Law, 2010; Melnik-Leroy et al., 2022; Zhang & Peng, 2017) the results of this study are more aligned with previous studies that reported mixed results, such as better perception for some sounds and better production for others (e.g. Hao & de Jong, 2016) or which report a lack of an overall significant correlation between perception and production data (e.g. Fabra & Romero, 2012; Kartushina & Frauenfelder, 2014; Peperkamp & Bouchon, 2011). Based on this finding, the results of the current study seem to fit better to the assumptions of the revised model (SLM-r) that "production and perception coevolve without precedence" (Flege & Bohn, 2021, p. 64).

Finally, orthographic cues did not seem to affect these participants' productions to a large extent overall, contrary to predictions that the differences in orthographic patterns between the two languages would affect vowel production based on previous research findings (e.g. Bassetti, 2017; Bassetti et al., 2015; Koda, 1989; Nimz & Khattab, 2020; Stoehr & Martin, 2022). Spectral analyses showed a large degree of overlap between the vowels produced in each task and context, although more vowels were closer to NE productions in the elicitation task, suggesting that participants may have benefited from listening to and imitating NE speakers in their productions rather than reading the orthographic forms. However, the effect of task on intelligibility data only reached significance in three vowels, two of which were rated as more intelligible in the wordlist-reading task and one as more intelligible in the elicitation task. The effect of task was mostly evident in the duration of vowels, most of which were significantly longer in the wordlist-reading task. However, this was expected due to the nature of the task, and did not affect intelligibility ratings, since NE raters did not appear to rely extensively on duration as a cue to distinguish between the members of contrastive pairs.

Contrary to predictions, overall perception was also not significantly affected by context in the present study. This is in line with previous findings for more experienced learners, who were found to disregard contextual variation and perceive L2 vowels more consistently (Balas, 2018; Levy & Strange, 2008). However, strong contextual effects were observed in individual vowels, confirming previous studies (e.g. Bohn & Steinlen, 2003) that different vowels may be affected by context to differing degrees. In line with Carlet and Cebrian (2014) and Thomson and Derwing (2016), these findings could not be explained by word-frequency or word-familiarity effects; although the present study did not directly examine their effect, vowel identification scores did not reveal any clear patterns indicating an effect of lexical frequency or familiarity.

6. CONCLUSION

This study aimed to provide an in-depth examination of the perception and production of L2 English vowels by CYG learners, compare them to the production patterns of NE speakers, and evaluate the effects of the L1, as well as orthographic and contextual effects. The results clearly demonstrated the influence of the L1 on both the perception and production of L2 segments, supporting the assumptions of current models of speech perception and production.

In identifying the difficulties that learners might encounter and how their L1 can influence L2 perception and production, this study can guide EFL teachers as well as course developers and material writers who can take advantage of these findings to update and improve language pedagogy not only for CYG learners of English, but also for other groups of learners whose L1 involves a smaller vowel inventory than their target L2. More specifically, this study can assist in the development of appropriate EFL curricula that aim to provide learners with adequate opportunities to practice challenging L2 segments, as well as exposure to input that preserves phonological contrasts between challenging L2 phonemes. In this way, teachers can guide learners in developing intelligible speech by targeting problematic segments, enabling them to better recognize and produce them to achieve their communication goals.

INTERESTS STATEMENT

There are no known conflicts of interest to disclose.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Dimitra Dimitriou: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Data availability

Data uploaded on Mendeley Data, doi: 10.17632/8xmms9dp52.1 (under embargo until 9 Nov 2024).

ACKNOWLEDGMENTS

The author thanks Dr Antri Kanikli, Dr Daniel Waller and Dr Daniel Matthias Bürkle for their insightful comments and guidance on earlier drafts of this paper.

References

Aliaga-Garcia, C., Mora, J.C., 2009. Assessing the effects of phonetic training on L2 sound perception and production. In: Watkins, M.A., Rauber, A.S., Baptista, B.O. (Eds.), Recent Research in Second Language Phonetics/phonology: Perception and Production. Cambridge Scholars Publishing, pp. 2–31.
Arvaniti, A., 1999. Cypriot Greek. J. Int. Phon. Assoc. 29 (2), 173–178.
Arvaniti, A., 2010. A (brief) overview of the phonetics and phonology of Cypriot Greek. In: Voskos, A., Goutsos, D., Mozer, A. (Eds.), The Greek Language in Cyprus from Antiquity to Today. University of Athens, pp. 107–124.
Baker, W., Trofimovich, P., 2006. Perceptual paths to accurate production of L2 vowels: The role of individual differences. International Review of Applied Linguistics in Language Teaching 44 (3), 231–250.
Balas, A., 2018. English vowel perception by Polish advanced learners of English. Canadian Journal of Linguistics/revue Canadienne De Linguistique 63 (3), 309–338.
Barriuso, T.A., Hayes-Harb, R., 2018. High variability phonetic training as a bridge from research to practice. CATESOL Journal 30 (1), 177–194.
Bassetti, B., 2017. Orthography affects second language speech: Double letters and geminate production in English. J. Exp. Psychol. Learn. Mem. Cogn. 43 (11), 1835–1842.
Bassetti, B., Escudero, P., Hayes-Harb, R., 2015. Second language phonology at the interface between acoustic and orthographic input. Appl. Psycholinguist. 36 (1), 1–6.
Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 (1), 1–48.
Best, C.T., 1995. A direct realist view of cross-language speech perception: New directions in research and theory. In: Strange, W. (Ed.), Speech Perception and Linguistic Experience: Issues in Cross-Language Research. York Press, pp. 171–204.
Best, C.T., Tyler, M., 2007. Non-native and second-language speech perception: Commonalities and complementarities. In: Bohn, O.S., Munro, M.J. (Eds.), Language Experience in Second Language Speech Learning: in Honor of James Emil Flege. John Benjamins Publishing Company, pp. 13–34.
Bohn, O.S., Bundgaard-Nielsen, R.L., 2008. Second language speech learning with diverse inputs. In: Piske, T., Young-Scholten, M. (Eds.), Input Matters in SLA. Multilingual Matters, pp. 207–218.
Bohn, O.S., Flege, J., 1997. Perception and production of a new vowel category by adult second language learners. In: James, A., Leather, J. (Eds.), Second-Language Speech: Structure and Process. Mouton de Gruyter, pp. 53–74.

Bohn, O., Steinlen, A.K., 2003. Consonantal context affects cross-language perception of vowels. In: Proceedings of the 15Th International Congress of Phonetic Sciences, pp. 2289–2292.
Carlet, A., and Cebrian, J. 2014. Training Catalan speakers to identify L2 consonants and vowels: A short-term high variability training study. In: Proceedings of the 7th International Symposium on the Acquisition of Second Language Speech, 5, 85-98.
Casillas, J.V., 2019. Phonetic category formation is perceptually driven during the early stages of adult L2 development. Lang. Speech 63 (3), 550–581.
Cebrian, J., 2006. Experience and the use of duration in the categorization of L2 vowels. J. Phon. 34 (3), 372–387.
Cebrian, J., 2007. Old sounds in new contrasts: L2 production of the English tense-lax vowel distinction. Proceedings of the 16Th International Congress of Phonetic Sciences.
Cebrian, J., 2019. Perceptual assimilation of British English vowels to Spanish monophthongs and diphthongs. J. Acoust. Soc. Am. 145 (1), EL52-EL58.
Cebrian, J., Carlet, A., Gorba, A., and Gavaldà, N. 2019. Perceptual training affects L2 perception but not cross-linguistic similarity. In: Proceedings of the 19th International Congress of Phonetic Sciences, Canberra, Australia, 929-933.
Coutsougera, P., 2007. The impact of orthography on the acquisition of L2 phonology: Inferring the wrong phonology from print. Proceedings of the Phonetics Teaching and Learning Conference.
Deterding, D., 1997. The formants of monophthong vowels in Standard Southern British English pronunciation. J. Int. Phon. Assoc. 27 (1–2), 47–55.
Di Paolo, M., Yaeger-Dror, M., Wassink, A.B., 2011. Analyzing vowels. In: Di Paolo, M., Yaeger-Dror, M. (Eds.), Sociophonetics: A Student's Guide. Routledge, pp. 87–106.
Dimitriou, D., 2019. L2 acquisition and production of the English rhotic by L1 Greek-Cypriot speakers: The effect of L1 articulatory routines and phonetic context. Philologia 16 (16), 45–64.
Dimitriou, D., 2022. Perceptual training effects on the perception and production of L2 English vowels by L1 Greek-Cypriot learners. In: Kkese, E. (Ed.), Speech perception and production in L2. Cambridge Scholars Publishing, pp. 125–151.
Fabra, L.R., Romero, J., 2012. Native Catalan learners' perception and production of English vowels. J. Phon. 40 (3), 491–508.
Flege, J.E., 1995. Second language speech learning: Theory, findings and problems. In: Strange, W. (Ed.), Speech Perception and Linguistic Experience: Issues in Cross-Language Research. York Press, pp. 233–277.
Flege, J.E., Bohn, O., 2021. The revised speech learning model (SLM-r). In: Wayland, R. (Ed.), Second Language Speech Learning: Theoretical and Empirical Progress. Cambridge University Press, pp. 3–83.
Flege, J.E., Wayland, R., 2019. The role of input in native Spanish late learners' production and perception of English phonetic segments. Journal of Second Language Studies 2 (1), 1–44.
Georgiou, G.P., 2019. Bit and beat are heard as the same: Mapping the vowel perceptual patterns of Greek-English bilingual children. Lang. Sci. 72, 1–12.
Hao, Y., de Jong, K., 2016. Imitation of second language sounds in relation to L2 perception and production. J. Phon. 54, 151–168.
Hay, J., Warren, P., Drager, K., 2006. Factors influencing speech perception in the context of a merger-in-progress. J. Phon. 34 (4), 458–484.
Huensch, A., Tremblay, A., 2015. Effects of perceptual phonetic training on the perception and production of second language syllable structure. J. Phon. 52, 105–120.
Hughes, A., Trudgill, P., Watt, D., 2013. English accents and dialects: An introduction to social and regional varieties of English in the British Isles. Routledge.
Hutchinson, A.E., Dmitrieva, O., 2022. Exposure to speech via foreign film and its effects on non-native vowel production and perception. J. Phon. 95,. https://doi.org/10.1016/j.wocn.2022.101189 101189.
Iverson, P., Evans, B.G., 2007. Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. J. Acoust. Soc. Am. 122 (5), 2842–2854.
Iverson, P., Evans, B.G., 2009. Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. J. Acoust. Soc. Am. 126 (2), 866–877.
Iverson, P., Kuhl, P.K., Akahane-Yamada, R., Diesch, E., Kettermann, A., Siebert, C., 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. Cognition 87 (1), B47–B57.
Kartushina, N., Frauenfelder, U.H., 2014. On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. Front. Psychol. 5. https://doi.org/10.3389/fpsyg.2014.01246.
Kerswill, P., Watson, K., 2014. Phonological considerations in sociophonetics. In: Holmes, J., Hazen, K. (Eds.), Research Methods in Sociolinguistics: A Practical Guide. John Wiley & Sons, pp. 136–148.
Kluge, D.C., Rauber, A.S., Reis, M.S., and Bion, R.A.H. 2007. The relationship between the perception and production of English nasal codas by Brazilian learners of English. In: Proceedings of Interspeech 2007, Belgium, 2297-2300.
Koda, K., 1989. Effects of L1 orthographic representation on L2 phonological coding strategies. J. Psycholinguist. Res. 18, 201–222.
Koutsoudas, A., Koutsoudas, O., 1962. A contrastive analysis of the segmental phonemes of Greek and English. Lang. Learn. 12 (3), 211–230.
Kyprianou, M., 2015. Teaching and learning English pronunciation in Cyprus: Survey, analysis and challenges. University of Cyprus (Unpublished doctoral dissertation).

Lengeris, A., and Hazan, V. 2007. Cross-language perceptual assimilation and discrimination of southern British English vowels by Greek and Japanese learners of English. In: Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany, 3, 1641-1644.
Lengeris, A. 2009. Individual differences in second-language vowel learning [Doctoral dissertation, University College London]. UCL Discovery. https://discovery.ucl.ac.uk/id/eprint/19029/1/19029.pdf.
Lengeris, A. 2018. Computer-based auditory training improves second-language vowel production in spontaneous speech. The Journal of the Acoustical Society of America, 144(3), EL165–EL171.
Leung, K.K., Jongman, A., Wang, Y., Sereno, J.A., 2016. Acoustic characteristics of clearly spoken English tense and lax vowels. J. Acoust. Soc. Am. 140 (1), 45–58.
Levy, E.S., Law, F.F., 2010. Production of French vowels by American-English learners of French: Language experience, consonantal context, and the perception-production relationship. J. Acoust. Soc. Am. 128 (3), 1290–1305.
Levy, E.S., Strange, W., 2008. Perception of French vowels by American English adults with and without French language experience. J. Phon. 36 (1), 141–157.
Lobanov, B.M., 1971. Classification of Russian vowels spoken by different speakers. J. Acoust. Soc. Am. 49 (2), 606–608.
Mayr, R., Escudero, P., 2010. Explaining individual variation in L2 perception: Rounded vowels in English learners of German. Biling. Lang. Congn. 13 (3), 279–297.
Melnik-Leroy, G.A., Turnbull, R., Peperkamp, S., 2022. On the relationship between perception and production of L2 sounds: Evidence from anglophones' processing of the french/u/–/y/contrast. Second. Lang. Res. 38 (3), 581–605.
Mildner, V., 2013. Experimental and quasi-experimental research in clinical linguistics and phonetics. In: Müller, N., Ball, M. (Eds.), Research Methods in Clinical Linguistics and Phonetics: A Practical Guide. John Wiley & Sons, pp. 28–47.
Nagle, C., 2018. Perception, production, and perception-production: Research findings and implications for language pedagogy. Contact Magazine, 5–12 http://contact.teslontario.org/perception-production/.
Nimz, K., Khattab, G., 2020. On the role of orthography in L2 vowel production: The case of polish learners of German. Second. Lang. Res. 36 (4), 623–652.
Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., Trent-Brown, S.A., 2008. Acoustic and perceptual similarity of Japanese and American English vowels. J. Acoust. Soc. Am. 124 (1), 576–588.
Papachristou, V., 2011. Explicit vs. implicit pronunciation teaching to Greek children: The case of the acquisition of English vowels. Selected Papers on Theoretical and Applied Linguistics 19, 371–381.
Peperkamp, S., and Bouchon, C. 2011. The relation between perception and production in L2 phonological processing. In: Proceedings of Interspeech 2011, Florence, Italy, 161-164.
Rato, A., Carlet, A., 2020. Second language perception of English vowels by Portuguese learners: The effect of stimulus type. Ilha Do Desterro 73 (3), 205–226.
Rauber, A., Rato, A., Kluge, D., and Santos, G. 2012. TP – Perception Tests / Perception Training Tasks (Version 3.1) [Computer software]. http://www.worken.com.br/tp/tp_install.html.
Sakai, M. 2016. (Dis)connecting perception and production: Training native speakers of spanish on the english /iː/-/ɪ/ distinction [Doctoral dissertation, Georgetown University, Washington]. Georgetown University ProQuest Dissertations Publishing.
Simaki, V., Mporas, I., Kondyli, M., 2015. Acoustic Differences between Standard Greek and Cypriot Greek. In: Proceeding of the 17th International Conference on Speech and Computer, University of Patras. University of Patras, pp. 71–78.
Stoehr, A., Martin, C.D., 2022. The impact of orthographic forms on speech production and perception: An artificial vowel-learning study. J. Phon. 94,. https://doi.org/10.1016/j.wocn.2022.101180 101180.
Suzukida, Y., Saito, K., 2019. Which segmental features matter for successful L2 comprehensibility? Revisiting and generalizing the pedagogical value of the functional load principle. Lang. Teach. Res. 25 (3), 431–450.
Themistocleous, C., 2017a. Dialect classification using vowel acoustic parameters. Speech Comm. 92, 13–22.
Themistocleous, C., 2017b. The nature of phonetic gradience across a dialect continuum: Evidence from Modern Greek vowels. Phonetica 74 (3), 157–172.
Themistocleous, C., Logotheti, A., 2016. Standard Modern Greek and Cypriot Greek vowels: A sociophonetic study. Modern Greek Dialects and Linguistics Theory 6 (1), 178–184.
Thomson, R.I., 2011. Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. CALICO J. 28 (3), 744–765.
Thomson, R.I., Derwing, T.M., 2016. Is phonemic training using nonsense or real words more effective? In: Levis, J., Le, H., Lucic, I., Simpson, E., Vo, S. (Eds.), Proceedings of the 7th pronunciation in second language learning and teaching conference. Iowa State University, pp. 88–97.
Tyler, M.D., 2019. PAM-L2 and phonological category acquisition in the foreign language classroom. In: Nyvad, A.M., Hejná, M., Højen, A., Jespersen, A.B., Sørensen, M.H. (Eds.), A Sound Approach to Language Matters – in Honor of Ocke-Schwen Bohn. Aarhus University, pp. 607–630.
Wells, J.C., 1982. Accents of English. Cambridge University Press.
Zhang, K., and Peng, G. 2017. The relationship between the perception and production of non-native tones. Proceedings of Interspeech 2017, Stockholm, Sweden, 1799-1803.