LITERA
Vol. 23 No. 1, March 2024
https://journal.uny.ac.id/index.php/litera/issue/view/2619
DOI: https://doi.org/10.21831/ltr.v23i1.67775

Analysing the acoustic evidence of consonant productions among students at the university level

Siti Nur'Aini*
Pascasarjana Universitas PGRI Semarang, Indonesia
*Corresponding Author; Email: sitinuraini@upgris.ac.id

ABSTRACT

Learners of English as a second language (L2) whose first language (L1) is Indonesian tend to struggle to produce aspirated consonants. This study investigates whether the difficulties came from the interference of the L1 for the production of these sounds: [k], [b], [d], [g] in the final positions, and [p] and [t] in the stressed syllable. This study involved two cohorts of English department university students with different levels of fluency in L2 speech production. The L2 learners were asked to pronounce 25 words from a textbook previously used to teach them. The L2 learners were exposed to a British English-speaking speech model, which became a benchmark for the L2 learners' pronunciation by asking them to imitate the pronunciation. Annotation of data was conducted twice by a second annotator to ensure the objectivity of the scores given to the L2 learners which was analysed using paired sample t-test. Findings suggest that the sounds with the lowest success rate of production were [p] in the stressed syllable, [k], and [g] in the final position. The production was unsuccessful because the L2 learners did not have phonological awareness of how the L2 consonant sounds were produced near-natively and were affected by their L1. The lack of awareness led to the failure to produce [p], [k], and [g] sounds because these sounds did not exist in their L1 and interference of the L1 was embed to the L2 speech production. The formant analysis results using PRAAT indicate that there is an improvement in the participants' pronunciation after exposure to the native speaker's speech sound. The implication of this research is paramount for L2 learners and lecturers in highlighting the importance of targeted instruction and intervention to address the challenges in speech production. Contrasting the phonetic features of L1 and L2 sounds helped the learners to defer interference in their L2 speech production. This study encourages continuous assessment of L2 learners to ensure that they maintain the consistency of speech production to sound near-native.

Keywords: Consonant sounds, L1 interference, phonetic training

Article history

Submitted:	*Accepted:*	*Published:*
22 November 2023	*18 March 2024*	*25 March 2024*

Citation (APA Style):
Nur'Aini, S. (2024). Analysing the acoustic evidence of consonant productions among students at the university level. *LITERA*, *23*(1), 76-90. https://doi.org/10.21831/ltr.v23i1.67775.

INTRODUCTION

English as a communication medium across the globe is a common knowledge that has been widely recognised. Its expansion has reached even the world's remotest areas, and the English language is nowadays considered to belong to all nations instead of to "native English speakers." Still, it is considered a language that can absorb different elements of cultures where it is spoken as the language continues to expand and be accepted as a currency of cultural diplomacy (Aydinli & Aydinli, 2024; Babaoğlu, 2021). The notion supports the idea that different pronunciations are acceptable in most parts of the world (Kirkpatrick, 2014; Taglialatela, 2021).

Nevertheless, errors are expected to occur in the L2 learners of English's speech production because native speakers of English have certain features in their pronunciation that are not present in the L1 (mother tongue) of non-native speakers (Wagner et al., 2021). One of the factors that cause pronunciation variation is L1 interference, which influences pronunciation (Jianping, 2014). It has been shown in many studies that distinct speech sounds are preferred in human languages because they are easier to acquire. In English, the contrast between voiced and voiceless sounds in the final position often poses difficulties for L2 (foreign language) learners whose language does have a feature of aspiration. English words with similar endings, such as *log* and *lock*, will cause some pronunciation far from sounding near-native in L2 learners as the difference between the two words is subtle. These words have

aspiration, something that is not present in some L2 pronunciations. Aspiration is the strong burst of breath at the release of voiceless obstruents (Ruch & Harrington, 2014). English is one of the languages that have aspiration in their sound productions. It can be challenging when the pronunciation is produced by someone whose L1 does not have aspiration, like Indonesian learners. The issues above become the main focus of this article, which will explore the production of consonants in the stressed syllables and aspirated sounds at the end of the syllable.

Existing studies on speech production have been found in the work of Ader and Miljan (2015), who found the external factors of L1 Estonian learners in their L2 English. The evidence showed that the younger generation of Estonians had less interference in their speech because they had greater exposure to L2 English, thus enabling them to produce more native-like sounds. Noviyenty and Putri (2021) focused on vowel interference among senior students, which resulted in a conclusion that L1 Spanish affected the vowel pronunciation of L2 English. Similar account studies were done on Indonesian students' pronunciation (Qomariana et al., 2019) and Arab and Chinese students' speech production (Li, 2016; Thyab, 2016), all of which emphasise the interference of L1 in the L2 speech production. The importance of exposure to L2 and phonological training became the focus of Franklin & MacDaniel's (2016) study, where they discovered that L2 learners' pronunciation improved after the training.

This study aims to fill the gap in research by delving into the consonant variables, particularly in the words that have aspirated sounds. These sounds of [k], [b], [d], [g] in the final positions, and [p] and [t] in the stressed syllable were chosen because these sounds allicit pronunciation challenges among L1 Indonesian pronunciation. To capture the extent of interference and what potential strategies or interventions to attenuate the pronunciation difficulties, different teaching methods or pronunciation training programs to help L1 Indonesian speakers improve their L2 English pronunciation. To seek evidence of the interference of the L1 into the speech sound productions in Indonesian students, this paper presents two hypotheses: 1) H0: There is an interference of the L1 to L2 learners' pronunciation; and 2) H1: There is no interference of L1 to L2 learners' pronunciation. The first is the null hypothesis, which assumes that the mean of two paired samples was equal, and the second hypothesis is the alternative hypothesis, which assumes that the means of two paired samples were not equal. With the hypotheses presented above, this study focuses on two issues related to how the L2 learners produce aspirated sounds: [k], [b], [d], [g] in the final positions, and [p] and [t] in the stressed syllable. The choice of these consonants was grounded on the notion that those consonants exist in both L1 and L2. Those sounds will provide enough challenge for the L2 learners to pronounce and to test their ability to produce near-native sounds of the consonants, although it is expected that the production of those sounds might be different.

A foreign language learner will have the sound of L1 that appears in the pronunciation of the L2, a phenomenon called language interference, an unconscious transfer of the L1 to the L2 without meaning to do so due to the habit of using the L1 (Biglari & Struys, 2021; Liu et al., 2021). In their study, Liu et al. (2021) argued that cross-language interference happened because of the language system of the L1 speaker when producing the L2 and was not influenced by how others produced the language. The transfer could happen in written or spoken use of the language. In a more specific and practical context, Gass et al. (2020) argue that L1 affects L2 production because people have accents, dialects of the certain area where they are from, and the language acquisition process of the L2. Ellis (2016) stated that L1 interference happens because of the linguistic transfer of a learner's L1 into the L2 production. These arguments emphasise how the L1 affects L2 sound productions and that this is not considered negative. Two languages can have similar features, but mostly, they are different in phonemes, sounds, intonation, or syllable stress caused by the nature of the language itself. One of the influential factors in different speech productions is the L1 (Cox, 2015).

Studies on L1 interference in various contexts have been done in numerous cases, like the work of Hassanzadeh and Salehizadeh (2020); Zeng et al. (2023), who adopted the work of Macdonald et al. (1994), where they established the steps to improve speech production in L2 learners by assigning their participants to four activities: (a) traditional drilling activities, (b) self-study with tape recordings, (c) interactive activities, and (d) a no-intervention control condition. Hassanzadeh and Salehizadeh (2020) focused on the lexical stress of the speech and the importance of mechanical drilling and repetition to get the L2 learners correct pronunciation. This study indicated that the four activities were influential in shaping the way participants produce their speech.

Phonetics is especially essential in speaking skills, where speakers need to understand the features in their L1 or L2 languages to produce speech sounds to communicate or practice. Each language has linguistic features in speech production, such as syllable structure, word stress, and intonation, that affect how sounds are produced. This is why languages could be challenging to foreign language learners (Yuvayapan, 2019). Phonetics considers the study of how speech sounds are articulated, discerned, and decoded in written forms, and the importance of phonetic training has been investigated in a number of studies (Melnik & Peperkamp, 2021; Saito et al., 2022). In speaking a foreign language, pronunciation plays an important role in message delivery success. It is important to practice frequently to achieve near-native and comprehensible foreign language pronunciation. The works of many researchers show that phonetic training is essential in improving foreign language learners' pronunciation (Alzinaidi & Latif, 2019; Bakar & Ridhuan, 2015; Gilakjani, 2012). The studies focusing on the importance of training or exposure to foreign language speech productions can be found in the work of Saito et al. (2022) and Melnik and Peperkamp (2021), whose works focus on the high variability of phonetic training, while Arora et al. (2018) made a point on the importance of exposing foreign language audio to the learners of foreign language and train them how to pronounce the words like a native. The importance of modelling was asserted in the work of Lee-Kim (2016) and Bassetti et al. (2018) to help learners to imitate speech productions better, especially with the support of a controlled computer-assisted pronunciation tool (Tejedor-Garcia et al., 2020). Similar accounts of the importance of speech exposure and training were also found in the study by Cao (2016); Liakin et al. (2015) where they ensure that the participants get repeated exposure to L2 speech. These studies argued that different types of exposure and training were beneficial to improve foreign language learners' pronunciation.

Two of the most important elements in sound production are the knowledge of phonetics and phonology. Speech sounds are uniquely different in each language because of articulation, the strength of voice, tongue height, and placement (Cheng, 2021). These details might be easily skipped because of the phonological differences between languages, mainly because the phonologist focuses on whether the differences are contrastive or distinctive. For example, a speaker articulates the sound by complete or partial closure of their vocal tract to produce a consonant. The comparison between the aspirated [p] and the non-aspirated [p] can be seen in the production. Consonant [p^h] is pronounced by putting the lips together, letting the air build up behind the lips, and then releasing the puff of air out of the lips. The non-aspirated sound of [p] did not have the air build up behind the lips that the non-native L2 speakers usually produce.

Another example is the production of consonant [t], which is produced by raising the tongue until the tip touches the alveolar ridge; the sides of the tongue will also be up and touch the molars on the inside of the teeth, building up some air pressure and release the quiet sound of [t] as the air pressure is released. The slight differences are noticeable to the native speakers of L2, but the EFL learners can overlook the differences. When this phenomenon happens frequently, it results in the, unlike L2, production of the speech, albeit that it is still understandable. However, to achieve near-native speech production, it is encouraged that all speech sounds are produced in the correct places within the speech organs. Knowing the speech organs and how they work together to produce the near-native consonant sounds is important.

	Labial	Labiodental	Dental/Alveolar	Palatal	Velar	Glottal
Nasals	m		n	ɲ	ŋ	
Stops/Affricates	p b		t d	tʃ dʒ (ʃ)	k g	ʔ
Fricatives		(f) (v)	s			h
Approximants	w			y		
Laterals			L			
Rhotics			R			
( ) = present only in 'borrowed' words from another language						

Figure 1. Indonesian consonant chart

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Retroflex	Palatal	Velar	Uvular	Glottal
Plosive	p b			t d				k g		
Nasal	m			n				ŋ		
Tap or flap										
Fricative		f v	θ ð	s z	ʃ ʒ					h
Affricate				t͡s d͡z	t͡ʃ d͡ʒ					
Approximant				ɹ			j			
Lateral approximant				l						

Figure 2. English consonants chart

Figure 1 presents the symbol for consonant sounds that exist in the Indonesian language. Figure 2 presents the symbols for consonant sounds in the English language. They are comparable, and some sounds are produced similarly. However, some English consonants do not exist in the Indonesian language, causing difficulties for EFL learners in pronouncing the English consonants. These consonant sounds are the dental fricative sounds /ð/ and /θ/. Another difference is in the sound of /r/. In Indonesian consonant, this sound falls into the category of glide alveolar, with a strong roll of the tongue to produce the sound.

How words are produced is also different. Both languages have bilabial plosive sounds like /p/ and /b/. The difference lies in the aspiration of the sound, where Indonesian words are mostly not aspirated. This production results in a non-puffed /p/ sound when it is said in Indonesian.

Another feature of the Indonesian language is the occurrence of the word stress, which generally happens in the penultimate syllable of the word. Meanwhile, stress, intonation, and tone in English are essential parts of the language to prevent misunderstanding, miscommunication, and alteration of meaning. The difference in syllable stress between the Indonesian and English languages presents a challenge for EFL learners. To compare the language pair, a contrastive analysis was necessary to gain an overall understanding of the features (Haziri et al., 2023; Koptleuova et al., 2022). Both languages were contrasted against the background of similarities and provided input to applied disciplines such as foreign language and L1 speech productions. There are four steps in conducting contrastive analysis (Chen, 2020; Izquierdo & Blanco, 2020).

Adopting the four steps, the first step of language analysis was done by describing both languages in terms of their features, phonetic and speech productions, and phonological structure of the languages. The second step is selecting which parts of the language to analyse. Parts of language include consonants, vowels, diphthongs, word stress, phonemes, and many other aspects of language. A researcher should decide which aspect of the language they want to analyse by following the third step, contrasting the features. This stage includes choosing the same language features and comparing them to see how they differ. The last step is predicting how the speech is produced and what problems occur in the process. The common challenge in doing contrastive analysis is to set up the arbitrary of how speech should be produced (Razavi et al., 2016). To overcome this challenge, a benchmark that became the point of reference for the L2 learners in this study was defined. The benchmark is a British English-speaking person who pronounced the sample words taken from the Practice and Progress book (Alexander, 1980). The analysis is conducted using statistical software and sound analysis software called Praat, as elaborated in the following part.

The Praat system is a tool for ‘doing phonetics by computer’ (Boersma & Weenink, 2019). Many researchers have used it to analyse different sound formant types and modify the sounds to match the researcher’s needs (Magdin et al., 2019; Núñez Batalla et al., 2014; Rohani et al., 2019; Sai Vineeth et al., 2018). It works by using digital signal processing where the speech sound synthesis is happening. One of the advantages of using Praat is its ease of steps in terms of phonetic and phonological analysis in sound productions. The display is user-friendly and straightforward, enabling users to produce displays of speech data as it was produced in the vocal tracts. The display of vocal tract movement provides how sounds were produced and data for analysis. One of the benefits of using Praat for speech analysis is the improved accuracy in speech production because the speaker can see the shape of the

sound (Liu et al., 2020). Therefore, to achieve this purpose, this study uses Praat to help capture the speech sounds that enable the analysis process of comparing the speech sounds of the native L2 speaker and L2 learners (Magdin et al., 2019; Núñez Batalla et al., 2014; Sai Vineeth et al., 2018). The Praat system is used to provide a visual for the sounds produced by the L2 learners, which then are cross-referenced with the visual of sounds produced by the L2 native speaker. It helps the analysis because the charts and wavelength provide an image of where the sound production occurs. The sound image makes it easier to mark the differences, uniqueness, and interference of the L1 in the sounds produced by the L2 learners. Phonetic measurement and analysis using Praat software made it easier to see the distinct features of sounds produced by the L2 native speaker and by L2 learners. One of the essential aspects of speech production included in the analysis is finding the aspirated sounds in the L2 learners' pronunciation. To find the aspiration in the speech production between L2 native speakers and L2 learners, it is necessary to compare both languages in an analysis that compares the language features. To achieve that, the contrastive analysis is employed in this study, which is elaborated in the part that follows. This present study elicits L2 learners' speech production by taking it from the phonological awareness of the L2 learners. It aims to find whether the L2 learners can produce L2 aspirated sounds and to see if there was interference of L1 in their speech production.

METHOD

This study used a quasi-experimental design (Gopalan et al., 2020; Lacoste et al., 2021) to observe the comparison of speech production before and after phonetic training and exposure to an L2 native speaker model. The model was a female British English native speaker who read and recorded the words in Table 1 and Table 2. Her voice was used as the benchmark for the L2 learners to copy the pronunciation. Statistical analysis was conducted using a paired sample t-test (Geng & Jin, 2023).

Two cohorts of English department students of a teacher training university who were in the second semester of their studies were selected as the population with 87 participants. From the total number of participants, a simple random sampling was employed using this formula:

$$n/N \text{ or } 8.7/87 = 0.10$$

It gave every student in the population a 10% chance of being selected for analysis. The sample for this research was n/N or 8.7/87 = 0.10 (or 10%). Since it was impossible to take 8.7 samples, the participants were rounded up to nine L2 learners. These nine L2 learners were requested to produce the sounds in Table 1 and Table 2.

There were two instruments in this study. The first instrument was the book from where the words were taken, a textbook for Reading 1 subject, Practice and Progress, by Alexander (1980). The words were taken from Chapters 1-24. The book was chosen because it was used to teach the L2 learners in the first semester, and the words were taken from each chapter by picking every 10th word in the paragraphs and counting it forward. From all the words collected, the list was filtered to choose the ones that contain [k], [b], [d], [g] in the final positions and [p] and [t] in the stressed syllable. These consonants were chosen because they were believed to be challenging for English L2 learners. Below are the lists of filtered words from the book that the L2 learners had to pronounce:

Table 1. [k], [b], [d], and [g] in the final positions

k	b	d	g
Think	Cub	Code	Big
Leak	Robe	Send	Bag
Black	Grab	Bed	Bug
Bark		Broad	Dog
Cork			

Table 2. [p] and [t] in the stressed syllable

P	t
Pool	Attend
Pulled	Attack
Popular	Torn
Power	Competition
	Time

The second instrument was Theresa Bynes's (pseudonym) voice, a female British English native speaker who read and recorded the words in Tables 1 and 2. The way the native speaker read the words was used as a benchmark for L2 learners' pronunciation. The L2 learners had to record their voices by imitating the model. The reason behind this choice of this British English speaker was because the book was published by a British publisher and had British English spelling; it was best to invite a native speaker of British English to pronounce the words so that L2 learners could imitate the pronunciation by the model (Goswami, 2020). The recording of the model was then played in the classroom for the L2 learners to listen and repeat. This process was conducted throughout the end of the semester. Finally, the L2 learners were asked to record their pronunciation for the second time after they were exposed to the native speaker.

Data was gathered using recorded L2 learners' voices as the master data during the second semester of the L2 learners' study. The steps to collect data were administered to the L2 learners in three ways: (1) L2 learners recorded their speech productions for the words on the list in Table 1 and Table 2 before the exposure to speech sound benchmark and training; (2) the exposure of the speech sound benchmark and training; (3) L2 learners recorded their speech productions for the same words after the exposure to speech sound benchmark and training.

These three steps were critical to finding out if there was improvement or difference in speech production before and after the exposure. Initially, the L2 learners had little or did not have the skills and knowledge to produce near-native sounds. They pronounced the words however they could and recorded the sounds. The exposure stage involved a lecturer who showed the L2 learners where the speech organs were and how to use them to produce near-native sounds. The speech sounds produced by the British speaker were played repeatedly as the benchmark or model for the near-native sound productions during the training or exposure period of 12 weeks.

The L2 learners were often asked to practice the speech to ensure that near-native sounds were produced. Finally, the L2 learners were asked to produce and re-record the speech. The process of data gathering and analysis is displayed in Figure 4 below.

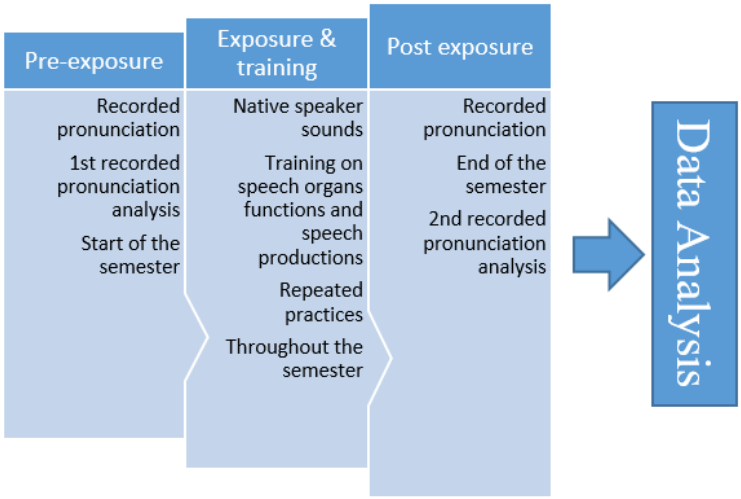


Figure 3. Data gathering process

The steps to analyse the data started with listening to the recording made by the L2 learners. The first set of audio to assess was the recording before L2 learners were exposed to the native speaker speech model. The second set of audio was that of the L2 learners after they were exposed to the native speaker's speech model. They were analysed the same way by giving a score to each word produced by the L2 learners. The score derived from a scale developed by Brown and Abeywickrama (2004, p. 157):

2.00: comprehensible, acceptable target form
1.00: comprehensible, partially near-native target form, and
0.00: silence or seriously far from near-native target form.

The three-scale rating helped assess the speech production and individually mark the participants' speech. After the assessment for both sets was complete, the average score the L2 learners earned was calculated.

Once the first annotation for the two sets of audio was completed, the second annotator did the same assessment of the audio using the same scoring steps and standards. This step was done to ensure objectivity in scoring the speech sounds produced by the L2 learners. The second annotator's role was to provide a different point-of-view on the speech production. The second annotator applied two assessments on the audio files before and after native-speaker exposure that the original annotator had a benchmark of assessment. With a linguistic background and experience in teaching Phonology at the university level, the second annotator had the qualification and experience to do the analysis.

Four groups of average scores were to be measured using the paired sample t-test, which will be divided into two pairs. The first pair was the average pre-exposure and post-exposure scores of nine L2 learners' pronunciation obtained from the first annotator. The next pair was the average pre and post-exposure scores obtained from the second annotator's assessment. The purpose of this test was to determine whether the exposure significantly influences the L2 learners' speech production. As a result, the first and second annotations by the two researchers found the four average scores as detailed in Table 3 below:

Table 3. Average scores for paired sample t-test

1st annotation		2nd annotation	
The pre-exposure average score of each participant	The exposure average score of each participant	The pre-exposure average score of each participant	The exposure average score of each participant

The steps to do Praat analysis are done by using the formant aspect of analysis where the concentration of acoustic energy was compared between speakers to enable real-time imaging and articulatory movements to capture the constant productions by the participants. The filtered sounds were the ones that had sound frequencies above 100 Hz and below 5000 Hz with a standard pitch setting of 75-600 Hz. The Table displayed in the Findings represents the formant frequencies corresponding to the shapes of the resonator in the oral cavity, which in turn corresponds to the position of the tongue and other speech organs.

RESULTS AND DISCUSSION

The analysis results are presented where each participant's average pre-exposure and post-score was compared to the average post-exposure score in the first annotation displayed in Table 4. The significance value for the t-test is <.001, which means the result is statistically significant, and there were significant differences in the score to prove that the exposure given for 12 weeks had good results.

Table 4: Paired t-test results from the 1st Annotation

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	preexposure - postexposure	-.36667	.21794	.07265	-.53419	-.19914	-5.047	8	<,001

The result of the analysis by the second annotator was done by calculating the average pre-exposure and post-exposure scores of the L2 learners. Table 5 displays the analysis results from the second annotator and indicates a significant value of <.001, which means that the results are statistically significant. The training and exposure provided good training grounds for the L2 learners because they could produce near-native speech sounds after the exposure.

Table 5. Paired t-test results from the second annotation

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Preexposure - Postexposure	-.17778	.10022	.03341	-.25482	-.10074	-5.322	8	<,001

Further evidence that there was an improvement in the L2 learners' pronunciation came from the calculation of the words pronounced by the L2 learners. The analysis result in the first annotation it appears that the ten hardest words to pronounce are shown in Table 6:

Table 6. List of words in Pre-exposure - first annotation with the least success rate

Words to pronounce	Score
Dog	0.20
Think	0.22
Big	0.40
Leak	0.40
Cub	0.60
Pool	0.70
Pulled	0.70
Bark	0.80
Black	0.80
Bug	0.80

The score was obtained by averaging the number assigned for each word produced by the nine L2 learners. Table 6 shows that seven out of the ten words ended with /g/ and /k/ in the final positions. They were the least accurately produced by the L2 learners before they were exposed to the native speaker sounds. It meant their pronunciation was silent or far from near-native target form. To compare the results of which words were the least successful to produce, the second annotation results were taken and displayed in Table 7:

Table 7. List of words in Pre-exposure - second annotation with the least success rate

Words to pronounce	Score
Think	0.33
Dog	0.44
Cub	0.67
Leak	0.67
Competition	0.67
Black	0.67
Robe	0.78
Bark	0.89
Pool	0.89
Cork	0.89

Table 7 displays the results from the second annotation before the exposure to the L2 native speaker sounds. Table 7 indicates that seven words are similar to the ones in Table 6. The words are [think, dog, cub, leak, black, robe, bark]. Five of the seven identical words had /g/ and /k/ sounds in the final position. The second annotator gave a slightly higher score for the words. Still, the similarities in judgment showed that the assessment indicated the words with /g/ and /k/ in the final position indeed were challenging for the L2 learners. One of the causes why the sounds/g/ and /k/ were the hardest to produce was because the L2 learners could not tell the difference between voiced and voiceless velar sounds. Not many Indonesian words end in /g/ sounds, and this absence has led to the unsuccessful reproduction of the sounds.

After exposure to the native-speaker sounds, L2 learners were asked to record their pronunciation, resulting in improved pronunciation, as shown in Table 8.

Table 8. List of words in post-exposure - first annotation with the least success rate

Words to pronounce	Score
Dog	0.70
Think	0.80
Bark	0.80
Cork	1.00
Leak	1.10
Black	1.10
Bag	1.10
Bed	1.10
Cub	1.20
Bug	1.20

The words that appear in Table 8 are almost similar; at least six words intersect with the words in Table 6. The difference was in the average score or success rate of the pronunciation. Table 8 indicates that there was an improvement in speech production. In the pre-exposure table, the word [dog] was scored at 0.20, which can be interpreted as the majority of L2 learners not being successful in producing the word. After the exposure, a significant increase in the score for the word [dog] indicates that more L2 learners were successful in producing near-native sounds of the word by imitating the model. The post-exposure analysis by the second annotator indicated a change in the word list that the L2 learners successfully produced.

Table 9. List of words in post-exposure - second annotation with the least success rate

Words to pronounce	Score
Attend	0.67
Bed	0.78
Think	0.89
Cub	0.89
Torn	0.89
Attack	0.89
Dog	1.00
Leak	1.00
Robe	1.00
Competition	1,11

The words in Table 9 that did not appear in the pre-exposure analysis by the second annotator now appear in the list. This indicates that the words were successfully produced before the exposure. However, it turned out that after the exposure, the L2 learners were not successful in copying the model and thus failed in producing the near-native pronunciation. The most probable cause was the misunderstanding about speech organ placement to produce the near-native pronunciation for the word [attend and bed], as they were scored the least. Words with /k/ and /g/ appeared in the list as well, although not as many as in the list scored by the first annotator. The appearance of the words with /k/ and /g/ indicated that the sounds remained a challenge for the L2 learners. However, after the exposure, they could produce the words and got a score of 1.00, which means that their pronunciation for those words ending in /k/ and /g/was comprehensible, partially near-native target forms.

Praat analysis results

The analysis result focuses on the formant of speech, where the concentration of acoustic energy around the frequency of the native speakers' speech waves was compared to the participants' speech formants. This part provides an excerpt of the results from one participant. The numbers extracted in Table 10 come from the formant frequencies of the 25 words pronounced by an L2 learner. These numbers are different from the ones in the statistical analysis because Table 10 displays frequencies that differentiate the formant of consonants. If there were two similar consonants, for example,/k/ and /g/ sounds, they could have close values. Table 10 indicates the length of time the L2 learner had taken to produce the sound and the frequency of the sounds.

Table 10. Formant of L2 learner's speech productions

WORDS	F1	F2
Think	650	945
Leak	677	1.832
Black	680	1.081
Bark	244	383
Cork	574	1.032
Cub	353	462
Robe	436	1.059
Grab	413	1.179
Code	744	1.092
Send	638	1.623
Bed	554	856
Broad	720	1.203
Big	439	880
Bag	899	1.116
Bug	704	1.647
Dog	744	1.092
Pool	903	1.413
Pulled	522	1.171
Popular	823	1.103
Power	563	994
Attend	629	1.492
Attack	940	1.201
Torn	781	1.762
Competition	674	1.674
Time	848	1.337

There are several observations to make. First, the formant frequencies differentiate the consonant relatively well. If two consonants are similar in one formant, they are differentiated in the other formant. For example, comparing [g] of 'bag' with [k] of 'attack', it can be seen that the first formant (F1) values are quite close, but the second formant (F2) values are clearly different. Similarly, if F2 values are relatively close, the F1 values are different. Therefore, calling formants the acoustic signatures of vowels is justified. The potential of formants to differentiate the consonants is to find out if the formant values connect between the articulatory features of the speech, especially with the position of speech organs to produce the sounds. The F1 and F2 are ordered based on the values to determine that the improvement has occurred in the pronunciation. The higher value indicates that the repeated intervention in terms of native speaker's sound exposure has yielded good results.

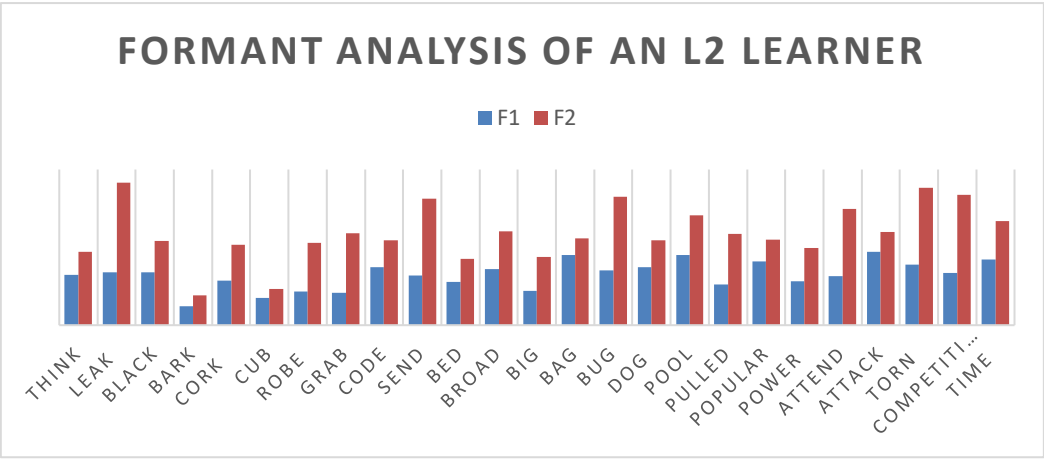
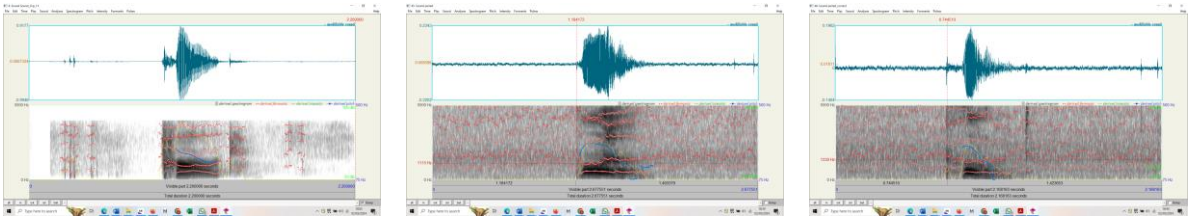


Figure 4. Formant analysis of an L2 learner

Figure 4 provides a comparison between the sounds produced by the native speaker and the L2 learner. The native speaker's pronunciation has a puff at the beginning of the speech, but the pre-

exposure sound produced by the L2 learners did not have that feature. Only after being exposed to training and practice was the participant able to produce the puff sound, indicated in the soundwave picture on the right column, which appears similar to the one produced by the L2 native speaker.



Native speaker's production | *Participant's attempt before exposure* | *Participant's attempt after exposure*

Figure 5. Waveform comparison for the word /pulled/

The excerpt of analysis using PRAAT software indicates a few stark differences in participants' pronunciation before and after the exposure to the native speaker's sound. Figure 5 shows the waveform of the native speaker's pronunciation of the word /pulled/ and compares it to the participant's attempt at producing the same sound before and after exposure to the native speaker's repeated training. The native speaker's pronunciation indicated a waveform where there was an increase of lines at the onset of the words and a faint but existing raised line at the end of the word. This increase in lines means that the /p/ sound was aspirated, and the consonant /d/ at the end of the syllable was released. The feature of the attempt by the L2 learners before they were exposed to the native speaker sound indicated the absence of both lines, meaning that the /p/ sound at the onset of the syllable was not aspirated, and the /d/ sound at the end was not released. When compared side by side like this, it was apparent that the L2 learners failed to produce the sound like the native speaker.

It had been established earlier that the null hypothesis agreed that there was an interference of the L1 to L2 learners' pronunciation, and the alternative hypothesis indicated that there was no interference of L1 to L2 learners' pronunciation. Given that all the findings and evidence have been presented, it becomes apparent that the speech productions of the L2 learners in the university context for the words given to them are influenced by the L1. Thus, the null hypothesis is accepted. Furthermore, all the results of the statistical analysis indicated a significant improvement in the L2 learners' pronunciation after they were exposed to the sound of the native speaker during the intense 12-week training. The most significant feature of the training that had the biggest impact on the pronunciation was the frequency and intensity, where they had to repeat many times to listen to the model and to reproduce the sounds. The prolonged exposure facilitates the development. Some of the participants indicated that they were successful in producing the sound by imitating the model. The success of rigorous training is consistent with the work of Deswarte et al. (2020) and Nishio and Joto (2022), who performed similar procedures to improve the pronunciation of the participants in their studies.

The role of the second annotator provided a balancing feature in this study, as they provided independent analysis. The unique thing about the results was the fact that although the sounds were annotated separately and independently by two annotators, the participants who consistently made errors were the same individuals. This indicated the presence of a strong influence of L1 and a lack of understanding of how speech organs work and how to use them properly to produce speech. The errors persist because these L2 learners had fossilisation (Ankerstein, 2017; Booth, 2015; Mutqiyyah & Muhammad, 2017; Nishio & Joto, 2022; Rahal, 2018) in their pronunciation, a phenomenon that frequently happens in L2 learners. Their difficulties in telling apart subtle consonant differences also led to incorrect speech production. Furthermore, the thing that caused the lack of success in producing the speech was the words were lack of context, as found in a study by Demirezen (2016), regardless of the fact that they had been taught using the book as the main learning material when during the training, the participants were presented with a list of words without context or explanation, while in real-life contexts, communication was rarely conducted in singular word pronunciation.

CONCLUSION

The assessment of the sounds produced by the L2 learners indicated that consonants /k/ and /g/ are the ones that give the most challenges to them. The two sounds are velar consonants, which should

be pronounced by putting the back part of the tongue against the soft palate, also known as the velum. The place was at the back part of the roof of the mouth. In this study, it was found that the voiced and voiceless velar sounds were the most difficult to produce because the Indonesian language has a devoicing process where a speech sound that should be voiced is made voiceless. The evidence indicated that the L2 learners could not distinguish between /k/ and /g/ sounds. This inability to distinguish the two sounds resulted in mispronunciation of the words ending in /k/ and /g/, such as *back, bag, and dog.* Several L2 learners pronounce the word *dog* as *dock.* It is apparent that mother tongue interference plays a part in the L2 learners' pronunciation, altering the English sounds to be slightly different from how they should be pronounced. Near native pronunciation was successfully produced by some L2 learners only after they were trained and exposed to the native speaker sounds. It is important to note that although the pronunciation was not similar to the native, the L2 learners' pronunciation was still understandable, and they did not alter the meaning of the words. Thus, their speech productions were acceptable, especially because their pronunciation indicated their identity as they carried their mother tongue in their speech. It offered language variety, and their deviance from the English sounds did not mean that they lacked the capabilities to imitate the sounds. It is just that the structure of the L2 was different from their L1 or mother tongue. This interference happens in many languages, and the Indonesian language happens to be one of them.

The implication from the research result indicate exposing them to the native speaker's voice and asking them to repeat the model speech sounds gave significant change in student's pronunciation, especially for difficult consonant sounds like [k], [b], [d], [g] in the final positions and [p] and [t] in the stressed syllable. The exposure was especially useful in evoking natural native-like pronunciation. The results from the targeted instruction and intervention to improve pronunciation indicated the necessity of exposure to the L2 sounds. Further investigation should seek whether continuous assessment could help maintain L2 learners' understanding of the phonetic features. It is encouraged that further research in this field is conducted for different consonant sounds that make Indonesian students find difficulties in pronouncing the words.

ACKNOWLEDGEMENT

The author would like to thank the participants of this study for providing essential data for analysis. Furthermore, thanks to the second annotator who did the second annotation and the native speaker whose voice became the benchmark in this study. Your contributions are highly appreciated.

REFERENCES

Ader, K., & Miljan, M. (2015). External factors and the interference of L1 Estonian on L2 English pronunciation: An apparent-time study. *Eesti Rakenduslingvistika Ühingu aastaraamat*, *11*, 21-36.

Alexander, L. G. (1980). *Practice and progress: An integrated course for pre-intermediate students*. Longman.

Alzinaidi, M. H., & Latif, M. M. A. (2019). Diagnosing Saudi students' English consonant pronunciation errors. *Arab World English Journal*, *10*(4), 180-193. https://dx.doi.org/10.24093/awej/vol10no4.14.

Ankerstein, C. (2017). Germans are not aiming for a fossilized form of English: A response to Booth (2015). *33*(4), 30-32. https://doi.org/10.1017/S0266078417000281.

Arora, V., Lahiri, A., & Reetz, H. (2018). Phonological feature-based speech recognition system for pronunciation training in non-native language learning. *Journal of the Acoustical Society of America*, *143*(1), 98-108. https://doi.org/10.1121/1.5017834.

Aydinli, E., & Aydinli, J. (2024). Exposing linguistic imperialism: Why global IR has to be multilingual. *Review of International Studies*, 1-22. https://doi.org/https://doi.org/10.1017/S0260210523000700.

Babaoğlu, R. (2021). Linguistic imperialism or cultural diplomacy? British Council and the People's Houses in early Republican Turkey. *Bilig*(96), 255-282. https://doi.org/10.12995/bilig.9610.

Bakar, Z. A., & Ridhuan, M. (2015). Importance of correct pronunciation in spoken English: Dimension of second language learners' perspective. *Social Sciences & Humanities*, *23*(S), 143-158.

Biglari, N., & Struys, E. (2021). Native language interference in English L2 word recognition and word integration skills. *Theory and Practice in Language Studies*, *11*(1), 1-11. https://doi.org/10.17507/tpls.1101.01.

Boersma, P., & Weenink, D. (2019). *PRAAT: Doing phonetics by computer*. In (Version 6.1.13) http://www.praat.org/4.3 .

Booth, J. E. (2015). The fossilization of non-current English pronunciation in German EFL teaching. *31*(4), 15-20. https://doi.org/10.1017/S0266078415000346.

Brown, D. H., & Abeywickrama, P. (2004). *Language assessment: Principles and classroom practices* (2nd ed.). Pearson Education ESL.

Cao, R. (2016). Improving English pronunciation teaching and learning via speech corpora of learners with dialectal backgrounds. *International Journal of Emerging Technologies in Learning*, *11*(4), 90-94. https://doi.org/10.3991/ijet.v11i04.5461.

Chen, H. (2020). A contrastive analysis of classroom-based language assessments. *English Language Teaching*, *13*(5), 110-114. https://doi.org/10.5539/elt.v13n5p110

Cheng, A. (2021). Maintenance of phonetic and phonological distance in the English and Korean back vowel contours of heritage bilinguals. *Journal of Phonetics*, *89*, 101109. https://doi.org/https://doi.org/10.1016/j.wocn.2021.101109.

Cox, R. (2015). 'What language are you?' A glimpse into multilingual childhoods. *English in Australia (0155-2147)*, *50*(1), 49-54. https://search.informit.org/doi/10.3316/informit.249015158310316

Demirezen, M. (2016). Assimilation as a co-articulation producer in words and pronunciation problems for Turkish English teachers. *16*(2), 477-509. https://doi.org/10.12738/estp.2016.2.0235.

Deswarte, C., Reyes-Rincón, J. H., Baquero, F. J., & Plata-PeñaFort, C. (2020). Fossilizing mistakes of French pronunciation in Spanish-speaking students. *12*(25), 59-76. https://doi.org/10.11144/Javeriana.m12-25.efpf.

Ellis, E. (2016). *The plurilingual TESOL teacher: The hidden languaged lives of TESOL teachers and why they matter*. De Gruyter Mouton.

Franklin, A., & McDaniel, L. (2016). Exploring a phonological process approach to adult pronunciation training. *American Journal of Speech-Language Pathology*, *25*(2), 172-182. https://doi.org/10.1044/2015_AJSLP-14-0172.

Gass, S. M., Behney, J., & Plonsky, L. (2020). *Second Language Acquisition: An introductory course*. Routledge. https://doi.org/10.4324/9781315181752

Geng, Y., & Jin, L. (2023). Anxiety and enjoyment of older learners of English in Chinese Universities of the third age. *Educational Gerontology*, *49*(6), 507-526. https://doi.org/10.1080/03601277.2022.2127699.

Gilakjani, A. P. (2012). The significance of pronunciation in English language teaching. *English Language Teaching*, *5*(4). https://doi.org/10.5539/elt.v5n4p96.

Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education*, *44*(1), 218-243. https://doi.org/10.3102/0091732X20903302

Goswami, A. (2020). Changing contours: The interference of the mother tongue on English speaking sylheti Bengali. *Journal of English as an International Language*, *15*(1), 100-134.

Hassanzadeh, M., & Salehizadeh, M. J. (2020). Focus on form options in second language pronunciation instruction: The case of lexical stress. *TESOL Journal*, *11*(2), e486. https://doi.org/10.1002/tesj.486.

Haziri, S., Dulaj, F., Neziri, S., & Duraj, P. (2023). A contrastive analysis of the definition of parts of speech in English and Albanian. *13*(4), 225-235. https://doi.org/10.36941/jesr-2023-0104.

Izquierdo, M., & Blanco, M. P. (2020). A multi-level contrastive analysis of promotional strategies in specialised discourse. *English for Specific Purposes*, *58*, 43-57. https://doi.org/10.1016/j.esp.2019.12.002

Jianping, L. (2014). A study of mother tongue interference in pronunciation of college English learning in China. *Theory & Practice in Language Studies*, *4*(8), 1702-1706. https://doi.org/10.4304/tpls.4.8.1702-1706

Kirkpatrick, A. (2014). English in Southeast Asia: Pedagogical and policy implications. *World Englishes*, *33*(4), 426-438. https://doi.org/10.1111/weng.12105

Koptleuova, K., Khairzhanova, A., Jumagaliyeva, U., Baiseuova, G., & Kurmangalieva, A. (2022). Contrastive analysis of cross-linguistic interference of trilingual oil workers. *Eurasian Journal of Applied Linguistics*, *8*(1), 13-27. https://doi.org/10.32601/ejal.911516

Lacoste, Y., Dancause, K., Bernard, P., & Gadais, T. (2021). A quasi-experimental study of the effects of an outdoor learning program on physical activity patterns of children with a migrant background: The PASE study. *Physical Activity and Health*, *5*(1), 236-249. https://doi.org/10.5334/paah.133

Li, F. (2016). Contrastive study between pronunciation Chinese L1 and English L2 from the perspective of interference based on observations in genuine teaching contexts. *English Language Teaching*, *9*(10), 90-100. http://dx.doi.org/10.5539/elt.v9n10p90

Liakin, D., Cardoso, W., & Liakina, N. (2015). Learning L2 pronunciation with a mobile speech recognizer: French/y/. *CALICO Journal*, *32*(1), 1-25. https://doi.org/10.1558/cj.v32i1.25962

Liu, H., Li, W., de Bruin, A., & He, Y. (2021). Should I focus on self-language actions or should I follow others? Cross-language interference effects in voluntary and cued language switching. *Acta psychologica*, *216*, 103308. https://doi.org/10.1016/j.actpsy.2021.103308.

Liu, X., Wu, D., Ye, Y., Xu, M., Jiao, J., & Lin, W. (2020). Improving accuracy in imitating and reading aloud via speech visualization technology. *International Journal of Emerging Technologies in Learning*, *15*(8), 144-160. https://doi.org/10.3991/IJET.V15I08.11475

Macdonald, D., Yule, G., & Powers, M. (1994). Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning*, *44*(1), 75-100.

Magdin, M., Sulka, T., Tomanová, J., & Vozár, M. (2019). Voice analysis using PRAAT software and classification of user emotional state. *International Journal of Interactive Multimedia and Artificial Intelligence*, *5*(6), 33-42. https://doi.org/10.9781/ijimai.2019.03.004

Melnik, G. A., & Peperkamp, S. (2021). High-variability phonetic training enhances second language lexical processing: Evidence from online training of French learners of English. *Bilingualism: Language and Cognition*, *24*(3), 497-506. https://doi.org/https://doi.org/10.1109/ELECSYM.2016.7861055

Mutqiyyah, R., & Muhammad, A. F. (2017). Developing mobile app of English pronunciation test using android studio. International Electronics Symposium, Denpasar, Indonesia.

Nishio, Y., & Joto, A. (2022). Improving fossilized English pronunciation by simultaneously viewing a video footage of oneself on an ICT self-learning system. *64*, 249-286. https://doi.org/10.1515/9783110736120-010.

Noviyenty, L., & Putri, M. I. (2021). Mother tongue interference towards students' English pronunciation: A case study in IAIN Curup. International Conference on Educational Sciences and Teacher Profession (ICETeP 2020). https://doi.org/10.2991/assehr.k.210227.049.

Núñez Batalla, F., González Márquez, R., Peláez González, M. B., González Laborda, I., Fernández Fernández, M., & Morato Galán, M. (2014). Acoustic voice analysis using the Praat programme: Comparative study with the Dr. Speech programme. *Análisis acústico de la voz mediante el progama Praat: estudio comparativo con el programa Dr. Speech (Spanish; Castilian)*, *65*(3), 170-176. https://doi.org/10.1016/j.otoeng.2014.05.007

Qomariana, Y., Puspani, I. A. M., & Rahayuni, N. K. S. (2019). Mother tongue interference on EFL: The case of English department students in Udayana university. Proceedings of the 65th TEFLIN international conference, Denpasar, Bali.

Rahal, A. (2018). Phonetic fossilization: Is it a matter of perfection or intelligibility? In (pp. 244-265). https://doi.org/10.4018/978-1-5225-4009-0.ch012.

Razavi, M., Rasipuram, R., & Magimai-Doss, M. (2016). Acoustic data-driven grapheme-to-phoneme conversion in the probabilistic lexical modeling framework. *Speech Communication*, *80*, 1-21. https://doi.org/10.1016/j.specom.2016.03.003

Rohani, G., Wahyu, M., & Rahmadsyah, R. (2019). Errors in pronouncing English phonemes: A Praat analysis. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, *3*(1), 49-63. https://doi.org/10.30743/ll.v3i1.1216

Ruch, H., & Harrington, J. (2014). Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics*, *45*, 12-25. https://doi.org/10.1016/j.wocn.2014.02.009.

Sai Vineeth, K. S., Phaneendhra, V., & Prince, S. (2018). *Identification of vowel phonemes for speech correction using PRAAT scripting and SPPAS.* International Conference on Communication and Signal Processing, India. https://doi.org/10.1109/ICCSP.2018.8524273.

Saito, K., Hanzawa, K., Petrova, K., Kachlicka, M., Suzukida, Y., & Tierney, A. (2022). Incidental and multimodal high variability phonetic training: Potential, limits, and future directions. *Language Learning*, *72*(4), 1049-1091. https://doi.org/10.1111/lang.12503.

Taglialatela, A. (2021). Authenticity and awareness of english as a lingua franca in english language classrooms. *Glottodidactica*, *48*(1), 103-121. https://doi.org/10.14746/gl.2021.48.1.06

Tejedor-Garcia, C., Escudero-Mancebo, D., Camara-Arenas, E., Gonzalez-Ferreras, C., & Cardenoso-Payo, V. (2020). Assessing pronunciation improvement in students of English using a controlled computer-assisted pronunciation tool. *IEEE Transactions on Learning Technologies*, *13*(2), 269-282, Article 9034084. https://doi.org/10.1109/TLT.2020.2980261.

Thyab, R. A. (2016). Mother-tongue interference in the acquisition of English articles by L1 Arabic students. *Journal of Education and Practice*, *7*(3), 1-4.

Wagner, M. A., Broersma, M., McQueen, J. M., Dhaene, S., & Lemhöfer, K. (2021). Phonetic convergence to non-native speech: Acoustic and perceptual evidence. *Journal of Phonetics*, *88*, 101076. https://doi.org/https://doi.org/10.1016/j.wocn.2021.101076.

Yuvayapan, F. (2019). Translanguaging in EFL classrooms: Teachers' perceptions and practices. *Journal of Language and Linguistic Studies*(2), 678-694. https://doi.org/10.17263/jlls.586811.

Zeng, J., Ponce, A. R., & Li, Y. (2023). English linguistic neo-imperialism in the era of globalization: A conceptual viewpoint. *Frontiers in Psychology*, *14*, 1149471. https://doi.org/10.17263/jlls.586811.