INDONESIAN JOURNAL OF APPLIED LINGUISTICS
Vol. 12 No. 3, January 2023, pp. 719-738

Available online at:
https://ejournal.upi.edu/index.php/IJAL/article/view/47474

https://doi.org/10.17509/ijal.v12i3.47474

Production and perception of English consonants by Yemeni EFL learners

Najah Ahmed Khamis Bin Hadjah[1] and Mohd Hilmi Hamzah[2*]
[1]Department of English, Seiyun University, Seiyun, Yemen
[2]Applied Linguistics Unit, School of Languages, Civilisation & Philosophy, UUM College of Arts & Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia

ABSTRACT

Previous studies on the perception-production correlation focused mainly on transcriptions and native English speaker evaluations for production accuracy assessments; only a few included acoustic measurements. This study aims to investigate the production and perception of six English consonants by Yemeni EFL learners of English using an acoustic measurement of second-language production. It has three objectives: (1) To investigate the extent to which word position influences the accuracy of Yemeni EFL learners' production of the target consonants, (2) to assess the extent to which word position affects their perception, and (3) to investigate the relationship between the overall production and perception of the investigated sounds. A quantitative research method was employed for collecting data from six Yemeni EFL postgraduate students from Universiti Utara Malaysia (UUM). In the production tasks, the researchers asked the speakers to produce seventy-two words with the target consonants three times in isolation (Experiment 1) and the other three times in a carrier sentence (Experiment 2). The researchers recorded the speakers' productions with an iPhone and later evaluated them acoustically via Praat. In the perception test, an AXB experiment was conducted. The findings showed that word position significantly affected the production, yet not the perception of the target sounds. Moreover, an insignificant positive moderate correlation was revealed between the overall production and perception of the target consonants. The findings have implications for second-language speech as well as pronunciation instruction. Teachers may put more focus on specific sound environments that lead learners to struggle while producing/perceiving particular English sounds.

Keywords: Correlation; EFL learners; perception; production; word position

First Received:	**Revised:**	**Accepted:**
18 June 2022	26 November 2022	8 December 2022
Final Proof Received:		**Published:**
27 January 2023		31 January 2023

How to cite (in APA style):
Bin Hadjah, N.A.K. & Hamzah, M. H. (2023). Production and perception of English consonants by Yemeni EFL learners. *Indonesian Journal of Applied Linguistics*, *12*(3), 719-738. https://doi.org/10.17509/ijal.v12i3.47474

INTRODUCTION

Second-language phonology research includes not only speech production yet also covers another crucial area of study for comprehending interphonology, which is second-language (L2) speech perception (Chan, 2011). Speech production is a synonym of "sound generation", and it refers to the mechanism of uttering articulated phonemes or words. It is an activity that is part of a complicated physical system, and it is generated by the cooperation of the lungs, the glottis (including the vocal folds), as well as the articulation tract (mouth in addition to the nose cavity) (Docio-Fernandez & Garcia-Mateo, 2015).

However, "speech perception" is occasionally referring to speech perception at the sublexical level (Black, 2019; Stasenko et al., 2015) since the evidence for the motor theory of speech perception is confined to tasks including syllable discrimination, which uses speech perception units (sublexical) rather than full spoken words (lexical) or sentences (Black, 2019; Junker et al., 2020).

* Corresponding Author
Email: hilmihamzah@uum.edu.my

Speech perception is assessed using tasks like identifying simple speech sounds (Stasenko et al., 2015). Typically, speech perception can be evaluated by the performance of two activities: discriminating between speech phonemes as well as identifying speech-sound categories. Discrimination requires determining if two sounds are different or similar (e.g., /da:/-/ba:/ versus /ba:/-/ba:/). In contrast, identification requires identifying a stimulus like a particular single sound or another (e.g., /da:/ versus /ba:/) (Stasenko et al., 2015).

English-language production has not captivated the interest of many Yemeni researchers (Baagbah et al., 2016), yet a few studies have recently been done for investigating Yemeni EFL learners' difficulties when producing certain English consonants (e.g., Al Mafalees, 2020; Bin Hadjah & Hamzah, 2022; Bin Hadjah & Jupri, 2018; Hamzah, Bin Hadjah, et al., 2020). The researchers of this study have not found any previous study yet that explored the difficulty in perceiving these six target consonants by Yemeni EFL learners. Thus, based on the available literature and the researcher's experience in teaching English to Yemeni adult EFL learners, the current study intends to examine how Yemeni EFL learners of English produce and perceive six English consonants (i.e., /p/, /v/, /θ/, /ð/, /ʧ/, as well as /ʤ/).

In this study, "perception" basically indicates the listener's ability to discriminate between two phonemically contrastive sounds (i.e., /p/-/b/, /v/-/f/, /θ/-/t/, /ð/-/d/, /tʃ/-/ʃ/, and /dʒ/-/ʒ/). Each target consonant, as revealed in past studies, was paired with the possible sound that it might be substituted with by Yemeni EFL learners. Some Yemeni EFL speakers were found to incorrectly produce /p/ as /b/(Al Mafalees, 2020; Bin Hadjah & Hamzah, 2022; Hamzah & Bin Hadjah, n.d.); /v/ like /f/(Bin Hadjah & Hamzah, 2022; Hamzah & Bin Hadjah, n.d.); /t/ for /θ/ and /d/ in place of /ð/ (Bin Hadjah & Hamzah, 2022); /ʃ/ instead of /tʃ/(Al Mafalees, 2020; Bin Hadjah & Hamzah, 2022; Bin Hadjah & Jupri, 2018; Hamzah, Bin Hadjah, et al., 2020); and /dʒ/ as /ʒ/(Bin Hadjah & Hamzah, 2022; Hamzah, Bin Hadjah, et al., 2020).

The link between the first-language (L1) and second-language phoneme systems has an influence on learning. For instance, according to the Speech Learning Model (SLM) (Flege, 1995; Flege & Bohn 2021) (see for its newly updated version SLM-r), the more perceived differences there are between L2 and L1 sounds, the more likely learners would create target-like categories. If learners of L2 have the ability of identifying phonetic distinctions between sounds of L1 and L2 that are closest in phonetic space, a new phonetic category for an L2 sound would be established, and learning "new" instead of "similar" L2 segments will be more successful for learners.

Likewise, the Perceptual Assimilation Model of Best (1994, 1995) (PAM) estimates that the challenge in distinguishing non-native phonemic contrasts is anticipated based on the relation between the phoneme inventories of L2 and L1. Moreover, discrimination would be most challenging if two non-native sounds are assimilated equivalently well or poorly to the same category of L1, and best if two sounds are assimilated into two distinct categories of L1.

For second-language who are in the process of constructing a second-language system, assimilation patterns have been assumed to be determined not only by L1-L2 relations yet also by how contrasting phonemes of a second language are related to one another in the developing phonological space between L1 and L2 (PAM-L2; Best & Tyler, 2007). This indicates that discriminating contrasts wherein one sound is perceived as a good exemplar of a specific category of L1 is expected to be good, and a new category is unlikely to emerge. Discrimination will also be good in circumstances where both sounds of L2 are perceived as equivalent to the same sound of L1; however, one is perceived to be a better fit than the other, with a new category creation anticipated just for the deviating sound. However, when the two sounds of L2 are assimilated to the same category of L1 yet are perceived as equally poor or equally good examples of that category, the L2-contrast discrimination would be poor initially, making learning challenging. Finally, based on the relations between the phonological system of L1 and L2, learning may be quite easy where none of the L2 sounds is perceived as associated with a specific L1 sound (uncategorised sounds).

Few past studies have examined perceiving and producing English sounds by Arab speakers of English, including Evans and Alshangiti's (2018) study on perceiving and producing English vowels in addition to (only) perceiving English consonants. Their study provided some evidence that perception and production are linked; English vowels were produced more accurately by Saudi learners who perceived them better. Evans and Alshangiti assured that future research is still needed to fully understand the possibility of learning new perceptual categories. For this reason, Alzinaidi and Abdel Latif (2019), who explored some difficulties in English consonant pronunciation, made recommendations for future research to examine the production and perception of English consonants by Arabs who learn English as an L2 and explore whether there is a correlation between consonant perception and production.

Therefore, following the recommendations by Alzinaidi and Abdel Latif (2019), the primary objective of the current investigation is to assess the production and perception of six English consonants by Yemeni EFL learners whose L1 is the Arabic language while their L2 is English. Unfortunately, little has been known about how to perceive and produce the six target sounds of the present study by Yemeni EFL learners, as well as the extent to which

word position affects the production and perception of these six consonants. Thus, this study aims to fill these gaps by answering the following three questions:

1. To what extent does word position, in which /p/, /v/, /θ/, /ð/, /tʃ/, and /dʒ/ occur, influence the accuracy of producing these six English consonants by Yemeni EFL learners?
2. To what extent does word position affect the accuracy of perceiving the six target consonants by Yemeni EFL learners?
3. Is there any correlation between the production and perception of /p/, /v/, /θ/, /ð/, /tʃ/, and /dʒ/ by Yemeni EFL learners?

The study has three hypotheses. H_1 postulates that word position affects the production of the six target consonants. H_2 also proposes that word position influences the perception of the investigated sounds. Additionally, H_3 predicts a correlation between the overall production and perception of the target sounds.

EFL learners vary in their capability to produce/perceive English sounds in different word positions. The current-study results could have implications for both English language instruction and learning. Teachers may focus more on specific sound environments that lead learners to struggle while producing/perceiving particular English sounds.

Production and Perception of English Sounds

By reviewing the current literature on producing and perceiving English consonants by EFL/ESL learners, it was noticed that only a few studies investigated the production and perception of the six target sounds of the present study (e.g., Lengeris & Nicolaidis, 2016; Syed et al., 2017), which incorporated Greek and Pakistani EFL learners, respectively. Other researchers (e.g., Nurfitriani, 2019) examined the perception in addition to the production of the target consonants of this study -except /p/, which was not investigated in her research.

Some previous studies reported more difficulty producing English sounds, whereas others discovered more difficulty perceiving English sounds. For example, Lengeris and Nicolaidis (2016) analysed how Greek learners of English could identify and produce English consonants. It was shown that certain consonants were easy to be identified yet challenging to be produced (e.g., /k/), whereas other sounds were proven to be troublesome to identify yet easy to produce (e.g., /d/). Likewise, Sioson and Chang (2017) found that some fricative contrasts were found to be easier for Thai learners to produce and perceive (e.g., /f/-/v/), while some contrasts were hard to produce and perceive (e.g., /θ/ and /ð/); however, other contrasts (e.g., /z/ and /s/) were easier to perceive yet challenging to produce in different phonological environments. Besides that, Lee (2019) found that while the most difficult task faced by Korean EFL college students in the perception experiment was to distinguish the pair /f-v/, the pair /θ-ð/ was regarded as the most troublesome for the participants in the production task (see also Kadiri et al., 2020; Sulistyorini & Wibowo, 2021).

Prior research examining the link between L2 learners' production and perception abilities has mainly relied on transcriptions and native English speaker judgements to evaluate production accuracy (Culleton, 2021). Consequently, this is an acoustic study that examines the connection between perception and production by acoustically assessing the speakers' production of the investigated consonants using Praat software.

The Effect of Word Position on the Production and Perception of English Sounds

Several past studies investigated the word-position effect on either the production (e.g., Alzinaidi & Abdel Latif, 2019; Bin Hadjah & Hamzah, 2022; Bin Hadjah & Jupri, 2018; Emran & Anggani, 2017; Hamzah, Bin Hadjah, et al., 2020; Huwari, 2019) or perception of English sounds (e.g., Kelly, 2019). However, a few researchers focused on examining the word-position effect on both producing and perceiving English sounds in one study. For instance, Jevring (2015) found that, generally, sounds in word-initial position were the easiest to perceive by Swedish speakers, yet the easiest to produce were those occurring word-medially.

Another study by Maiunguwa (2015) explored how Hausa ESL learners could produce and perceive the English fricatives /θ/, /ð/, and /v/ in word-initial and word-final positions. Generally, perceiving the investigated sounds was slightly better in word-initial than in word-final position. The findings of the production test, however, showed that the position of /θ/, /ð/, and /v/ played a significant role in the production of these three sounds. Producing these three fricatives in word-final position was more troublesome as compared to word-initial position.

The Correlation Between the Production and Perception of English Sounds

Despite the SLM and PAM's claimed link between speech production and perception, empirical evidence for such a link is not consistent (Schmitz et al., 2018) and is still unclear (Baese-Berk, 2019). Some past studies examined the connection between the two capacities in L2 learning, revealing a variety of outcomes from strong relationships (e.g., Chao et al., 2019; Syed, 2011) to moderate (e.g., Berti et al., 2020; Hattori, 2010), weak (e.g., Kaewchum, 2018; Sioson & Chang, 2017), or no link between these two modalities (e.g., Cheng et al., 2021; Zhang et al., 2016).

Chao et al. (2019) aimed to investigate the link between vowel perception and vowel production variability. Overall, their research confirmed the

notion that the association between vowel perception and production is strong and bidirectional. One potential explanation of such findings, as elaborated by Chao et al. (2019), was that aside from auditory goals, the system of speech motor employs the categorical perceptual boundary between two adjacent vowels to identify errors in its auditory output and; therefore, restrict variabilities of a vowel. An alternate explanation provided by Chao et al. was that our perception is shaped by our productions; hence, the variability of vowel production determines how adjacent vowels are perceived categorically (see also Syed, 2011).

Remarkably, a small number of studies have discovered a moderate link between perception and production. For example, Lersveen (2018) evidenced a generally positive moderate correlation between producing and perceiving the investigated sounds by Norwegian learners. Likewise, Hattori's (2010) study indicated a significant moderate association between /r/-/l/ identification and production. Such a moderate link was found by Berti et al. (2020) only in the fricative class yet, not in sonorants or stops (see also Zhang et al., 2021).

A few past researchers demonstrated that the association between perceiving and producing English sounds was weak. For example, Kaewchum (2018) showed that the link between the overall perception and production of the investigated consonants was generally weak positive. He also found that the correlations between the production and perception of /p/, /k/, /ɡ/, /f/, /v/, /s/, /z/, and /ʃ/ were very weak positive; however, between producing and perceiving /t/, it was weakly positive; while, between the /l/ production and perception, it was weakly negative (see also Sakai & Moorman, 2018; Sioson & Chang, 2017; Zhang et al., 2021)

Nevertheless, other studies proved no correlation between the perception and production of English vowels (e.g., Cheng et al., 2021; Zhang et al., 2016) or consonants (e.g., Huensch, 2013; Pei, 2022; Seo & Lim, 2016) by EFL/ESL listeners. For example, Seo and Lim (2016) found that those with lower English proficiency exhibited greater accuracy in production than perception, denoting no link between the two modalities. Likewise, no association was discovered by Huensch (2013) between improvements in production and perception (in the sentence context).

METHOD

In the present investigation, a quantitative research design was employed. In this type of research, numbers are used to explain the outcomes (Creswell, 2012) and to calculate the frequency of errors (Binasfour, 2018), including errors in perception (e.g., confusion between /p/ and /b/; /v/ and /f/, etc.) and production (e.g., substituting /p/ with /b/; /f/ for /v/, etc.).

Participants

The sample was selected using a purposive sampling technique. In the current investigation, six EFL postgraduates from UUM (a Malaysian public university) participated. There were three males and three females. All participants are Yemenis who were taught English as a foreign language while speaking Arabic as their L1. They had never visited any English-speaking countries. Moreover, there were no indications of speech or hearing impairments among the participants.

Selection of the participants was on the basis of their achievement on the English Language Placement Test (ELPT), which is required of international students who do not fulfil the English language proficiency requirements set by UUM. ELPT has a maximum of Band 9. The participants who scored Band 6 were chosen for this study to ensure that all participants have a similar level of English language proficiency.

The researchers explained the aims of the study and the data-collection technique to the participants. Additionally, the participants received assurances that the research report would not reveal their identities. Furthermore, the researchers told them that their participation in the study would be risk-free.

Instruments

Three experiments were carried out to collect data from participants: two production tests and one perception task. The two production experiments occurred before the perception test to prevent the participants from guessing the study focus. In the production experiments, the speakers were asked to produce three lists of seventy-two words: three times in isolation (Experiment 1) and the other three times in a carrier sentence (Experiment 2). The carrier sentence *"I say (the target word) three times."* was adapted from Hamzah (2013). A three-time repetition of a carrier sentence was carried out by several past researchers (e.g., Oh, 2019). There were four words for each investigated sound in the initial- and final-word positions: two disyllabic words besides two monosyllabic words. Nonetheless, all words in word-medial position had two syllables to allow for more accurate evaluations of their acoustic features.

In the production tasks, the total number of words analysed was 2592: 6 (target sounds) × 4 (four words for each investigated sound in each of the three-word positions) × 3 (positions of a word) × 6 (3 times: in isolation; 3 times: inserted into a carrier sentence) × 6 (number of speakers) = 2592 words. The researchers did a deep analysis within a speaker for the sounds that he/she produced three times in isolation and the other three times in a carrier sentence, using Praat for identifying the patterns occurring in each target sound across all three-word positions. This number of words is sufficient as a larger number would take a long time and a great deal of effort in the analysis process. The perception

experiment included a total number of 324 stimuli: 54 stimuli × 6 listeners.

Instruments' Validity and Reliability

The instruments were piloted before collecting data to ensure their validity and reliability. A jury comprising five experts with adequate experience teaching English and linguistics validated the production and perception instruments. All experts were Malaysian from UUM, except for Expert 4 and Expert 5, who were Yemeni affiliated to Seiyun University, a public university in Yemen. Regarding the assurance of reliability, Nunnally (1978) proposed a minimum level of (.7) Cronbach Alpha values for an instrument to be reliable. Cronbach's Alpha was (.868) for the production instruments and (.859) for the perception one, which are both more than (.7). Therefore, the instruments are considered reliable.

Procedures

In the production task, each speaker was given around 15 minutes to read each list, with a total reading time of approximately one and a half hours. The researcher recorded all the readings using an iPhone (Oppo F7). All the recordings were carried out within one week to ensure the consistency of the study.

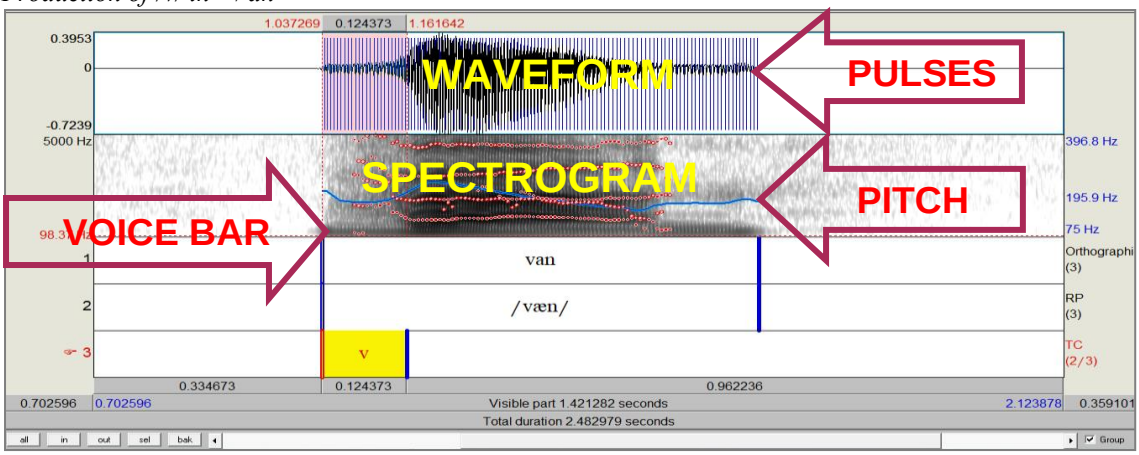
In the perception experiment, the AXB discrimination task (Best et al., 2001) was used. In this experiment, the participants listened to three stimuli for each trial. They were instructed in each trial to determine if X (the middle stimulus) was similar to A (the first stimulus) or B (the third stimulus). The participants' perception of the target sounds was examined using three vocalic contexts adapted from Kochetov (2004): /Cɑ/, /ɑCɑ/, and /ɑC/ (i.e., CV, VCV, and VC are nonsense words in which *C* represents the target consonant, while *V* indicates the vowel /ɑ/ used in the stimuli). The three vocalic contexts included the six investigated sounds across the three-word positions. Moreover, each target sound was paired with a possible alternative sound that EFL Arab learners may use to substitute. The stimuli were produced by a man who is a native speaker of Australian English.

Data Analysis

The following was the data analysis procedure: First, an iPhone was used to record the production experiment data, and then Praat was used to analyse it (Boersma & Weenink, 2022). Concerning the /v/ and /p/ analyses, the presence of the voice bar, pitch, and pulses in waveforms and spectrograms in Praat was visually inspected in order to determine whether these two target sounds were voiced or devoiced (see Figure 1). When one is present in waveforms or spectrograms, it indicates the presence of voicing, whereas if none of them exist, it suggests the absence of voicing (Hagiwara, 2009).

Figure 1

Production of /v/ in "Van"



In addition, /ð/ was easily distinct from /θ/ due to the presence of the voice bar, pulses, as well as pitch in the spectrograms and waveforms of /ð/, while /θ/ lacked all of these acoustic properties. Nevertheless, to determine whether /θ/ or /ð/ were stopped like either /t/ or /d/, this was accomplished by observing the existence of friction in the spectrogram of /θ/ and /ð/ and its disappearance in that of /t/ and /d/ (Firdaus et al., 2020). Regarding the two affricates, /tʃ/ and /dʒ/, these two sounds were distinguished on the spectrogram by a closure of the plosive portion followed by a sharp release of the fricative aperiodic noise portion (Alqarni, 2013).

For testing the extent to which word position in which the six target consonants occurred can affect the production and/or the perception of the six target consonants, the data of both production and perception were submitted to one-way variance analysis (ANOVA). Within and between-subjects ANOVAs were carried out to find out whether there could be any significant difference between the three-word positions when producing/perceiving the six

target consonants by the participants. The statistical significance test was conducted at Alpha =.05.

FINDINGS

The word-position effect on the accuracy of the production of the target English consonants by Yemeni EFL learners

With respect to the effect of word position on production, Levene's test was performed before conducting one-way ANOVA to determine whether the assumption of homogenous variances had been violated or not. The result revealed that this assumption was observed, $F(2,15) = .347$, $p = .712$. Therefore, ANOVA can be conducted.

ANOVA revealed that word position significantly affected the production of the target consonants at the p<.05 level, [$F(2,15) = 18.933$, $p = .000$] (see Table 1).

Table 1

ANOVA (Total Correct Production)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	8869.444	2	4434.722	18.933	.000*
Within Groups	3513.500	15	234.233		
Total	12382.944	17			

However, since there were three-word positions in this study, a multiple post-hoc comparison test was employed to figure out which of the three-word positions significantly differed from each other. Due to the small sampling size, a Bonferroni post-hoc test (see Table 2) was conducted (Pallant, 2020).

Table 2

Multiple Comparisons (Total Correct Production)

Multiple Comparisons

Dependent Variable: Total correct production

Bonferroni

(I) Word-position	(J) Word-position	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Initial	Medial	-17.500	8.836	.199	-41.30	6.30
	Final	35.833*	8.836	**.003***	12.03	59.64
Medial	Initial	17.500	8.836	.199	-6.30	41.30
	Final	53.333*	8.836	**.000***	29.53	77.14
Final	Initial	-35.833*	8.836	**.003***	-59.64	-12.03
	Medial	-53.333*	8.836	**.000***	-77.14	-29.53

*. The mean difference is significant at the 0.05 level.

Post hoc comparisons using a Bonferroni post-hoc test showed that the mean score for word-initial position (M =90.17, SD =19.104) significantly differed (p =.003) from word-final positions (M =54.33, SD =13.808). Similarly, there was a significant difference (p=.000) between the mean score for word-medial position (M =107.67, SD =12.127) and word-final position. However, there was no statistically significant difference (p =.199) between word-initial and word-medial positions.

As can be seen in Table 3, the production of the target sounds was greatly affected by word position.

Table 3

Percentages of Correct Production in the Three-Word Positions

Target sounds	Correct production in the three-word positions		
	Initial	Medial	Final
/p/	17%	26%	8%
/v/	27%	26%	3%
/θ/	25%	27%	28%
/ð/	26%	31%	11%
/ʧ/	20%	28%	25%
/ʤ/	11%	12%	2%

For instance, /p/, /v/, /ð/, and /ʤ/ were more difficult to produce when occurring word-finally compared to initial and medial word positions, whereas /θ/ and /ʧ/ were more challenging to produce in word-initial position in comparison to the other two-word positions. As a result, such an effect of word position greatly influenced the occurrence of substitutions of the investigated sounds across the three-word positions (see Appendix A).

The influence of word position on the accuracy of the perception of the target consonants by Yemeni EFL learners

Before carrying out the one-way ANOVA to test whether or not word position could significantly affect the perception of the six target consonants, the assumption of homogenous variances was checked using Levene's test to see if it had been violated. This assumption was detected, $F(2,15) = .273$, $p = .764$. However, ANOVA revealed that the effect of word position on the listeners' perception of the investigated consonants was insignificant at the $p<.05$ level, [$F(2,15) = 2.464$, $p = .119$] (see Table 4).

Table 4
ANOVA (Total Correct Perception)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	26.333	2	13.167	2.464	.119
Within Groups	80.167	15	5.344		
Total	106.500	17			

To summarise, word position was not found to significantly affect the perception of the target sounds. The findings, however, revealed that perceiving the investigated consonants was slightly better in word-medial position (M =15.83) in comparison to word-final (M =13.67) and word-initial positions (M =13.00).

This is clearly demonstrated in Table 5 that word position had a slight effect on perceiving the investigated sounds. As an example, /p/ and /ð/ were more problematic to perceive in word-final position, while /v/, /θ/, and /ʤ/ were harder to perceive in word-initial position. Regarding /ʧ/, fewer problems were found in its perception in word-medial position, yet more difficulty was exhibited in its perception when it occurred word-finally and word-initially.

Table 5
Percentages of Correct Perception in the Three-Word Positions

Target sounds	Percentages of correct perception in the three-word positions		
	Initial	Medial	Final
/p/	22%	33%	20%
/v/	20%	31%	28%
/θ/	28%	31%	30%
/ð/	30%	24%	22%
/ʧ/	24%	30%	24%
/ʤ/	20%	26%	28%

The correlation between the production and perception of /p/, /v/, /θ/, /ð/, /ʧ/, and /ʤ/ by Yemeni EFL learners

To assess the relation between producing and perceiving each target sound, a Pearson's r correlation coefficient was first computed. Second, the correlation between the overall production and perception of the six target consonants was also measured using the same test. For the interpretation of the results, Pearson's correlation coefficients by Dancey and Reidy (2007) (see Appendix B) were used.

First, concerning the correlation between producing and perceiving the English sound /p/, the results of the Pearson correlation test indicated that there was a significant positive strong correlation between the production and perception of /p/, [$r(.812) = 4$, $p < .050$] (see Table 6).

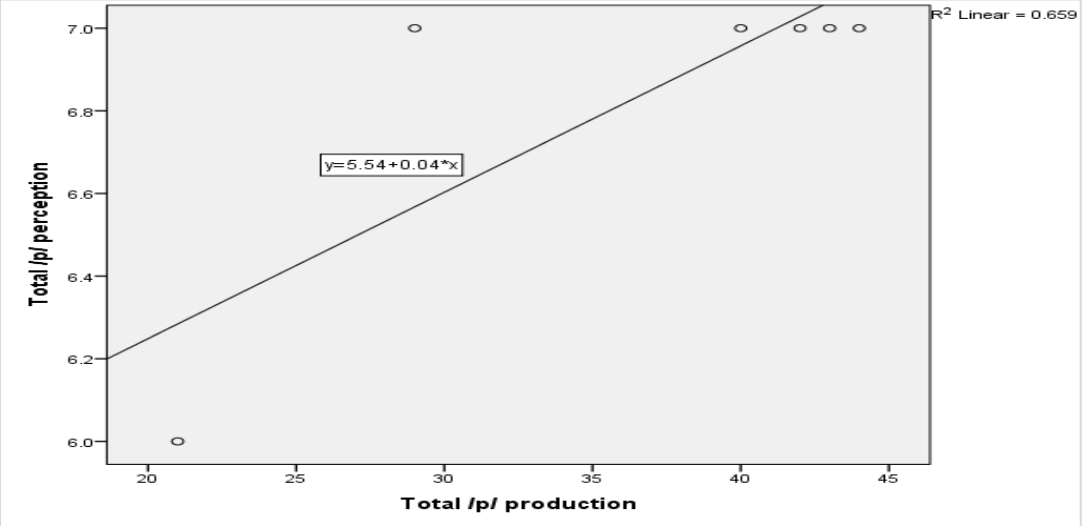
Table 6
The Correlation Between the Production and Perception of /p/

		Total /p/ perception
Total /p/ production	Pearson Correlation	.812
	Sig. (2-tailed)	.050*
	N	6

*. Correlation is significant at the 0.05 level (2-tailed).

A scatterplot summarises the findings (see Figure 2). Overall, there was a positive strong correlation between perceiving and producing /p/. Increases in /p/ production were correlated with increases in /p/ perception. For example, when the total correct production of /p/ was 21, the total correct perception was 6, and when the total correct production of /p/ was 29, the total correct perception was 7.

Figure 2
The Correlation Between the Production and Perception of /p/



Furthermore, the correlation between the production and perception of the voiceless affricate /ʧ/ was insignificant and positively strong, [$r(.750) = 4$, $p < .086$] (see Table 7).

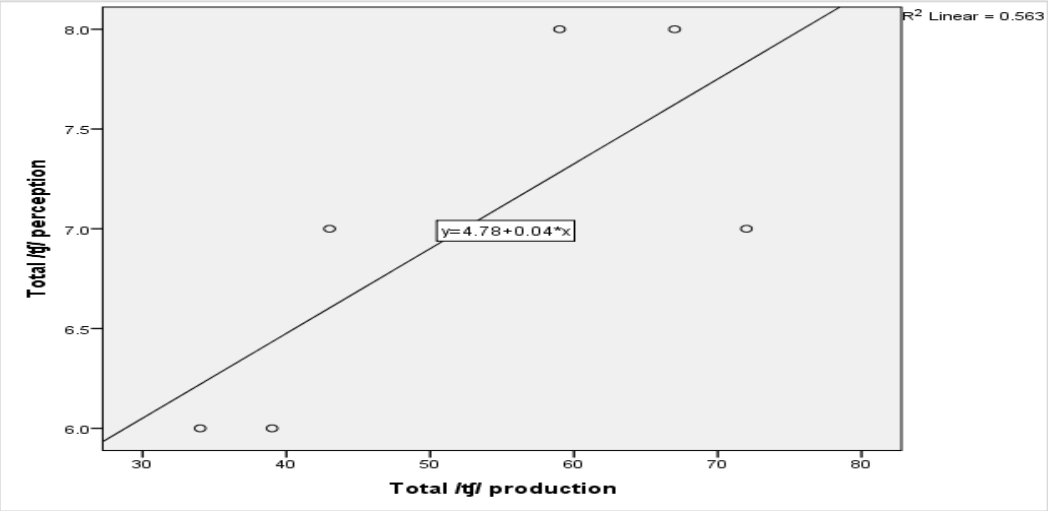
Table 7
The correlation between the production and perception of /ʧ/

		Total /ʧ/ perception
Total /ʧ/ production	Pearson Correlation	.750
	Sig. (2-tailed)	.086
	N	6

The scatterplot illustrates these results (see Figure 3). Generally, a positive strong correlation between perceiving and producing /ʧ/ was detected. Increases in the production of /ʧ/ were correlated with increases in /ʧ/ perception. As an example, when the total correct production of /ʧ/ was 34, the total correct perception was 6, and when the total correct production of /ʧ/ was 43, the total correct perception was 7.

Figure 3

The Correlation Between the Production and Perception of /ʧ/



As for the correlation between producing and perceiving the target consonant /v/, the Pearson correlation test showed an insignificant negative strong correlation between producing and perceiving /v/, [r(-.7) = 4, p < .153] (.660 can be approximated to .7) (see Table 8).

Table 8

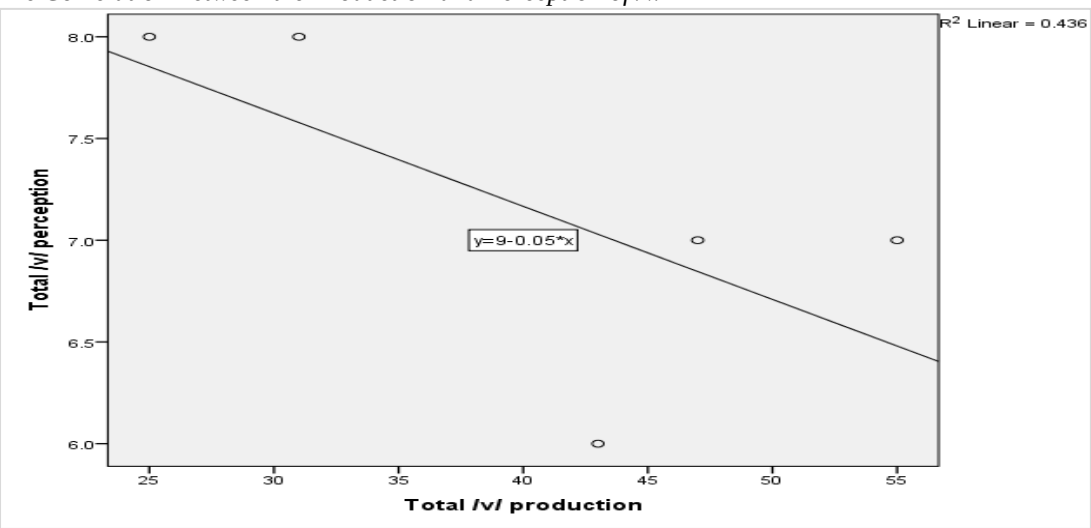
The Correlation Between the Production and Perception of /v/

		Total /v/ perception
Total /v/ production	Pearson Correlation	-.660
	Sig. (2-tailed)	.153
	N	6

The scatterplot provides a summary of the results (see Figure 4). In general, there was a negative (inverse) strong correlation between the production and perception of /v/. Increases in /v/ production were correlated with decreases in /v/ perception. For instance, when the total correct production of /v/ was 25, the total correct perception was 8. However, when the values of the total correct production of /v/ were 55, the total correct perception decreased to 7.

Figure 4

The Correlation Between the Production and Perception of /v/



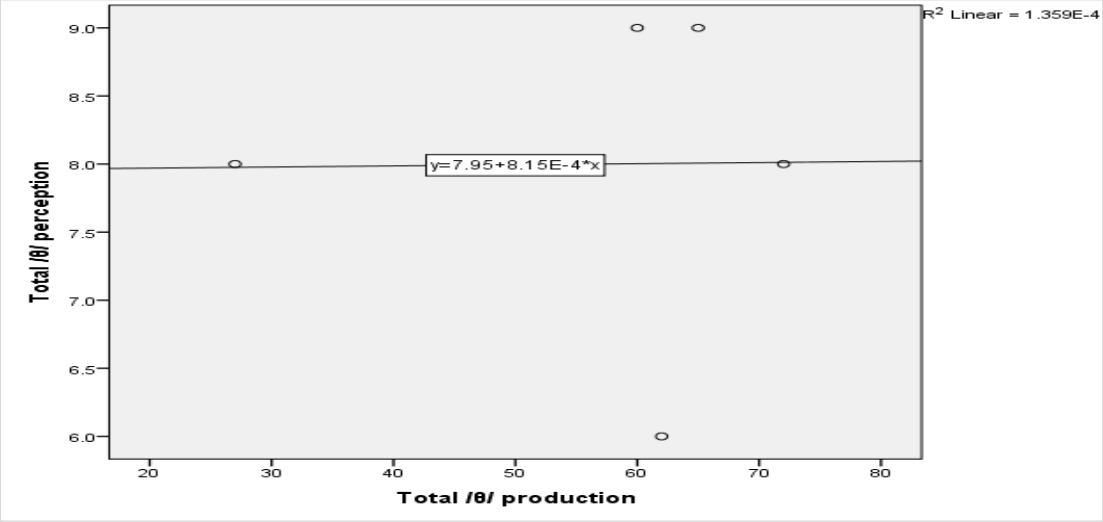
However, there was an insignificant positive weak correlation between the production and perception of /θ/, [r(.012) = 4, p < .983] (see Table 9).

Table 9
The Correlation Between the Production and Perception of /θ/

		Total /θ/ perception
Total /θ/ production	Pearson Correlation	.012
	Sig. (2-tailed)	.983
	N	6

A weak correlation indicates that when one variable decreases or increases, there is a lower likelihood of having a relationship with the second variable. As shown in Figure 5, when the production values change, the perception values do not change. It can be noticed in a weak correlation that the cloud is nearly flat or vertical. Therefore, the production and perception of /θ/ are weakly correlated since the scatter plot is almost flat.

Figure 5
The Correlation Between the Production and Perception of /θ/



Similarly, the correlation between the production and perception of the voiced affricate /ʤ/ was insignificant and positively weak, [$r(.234) = 4$, $p < .655$] (see Table 10).

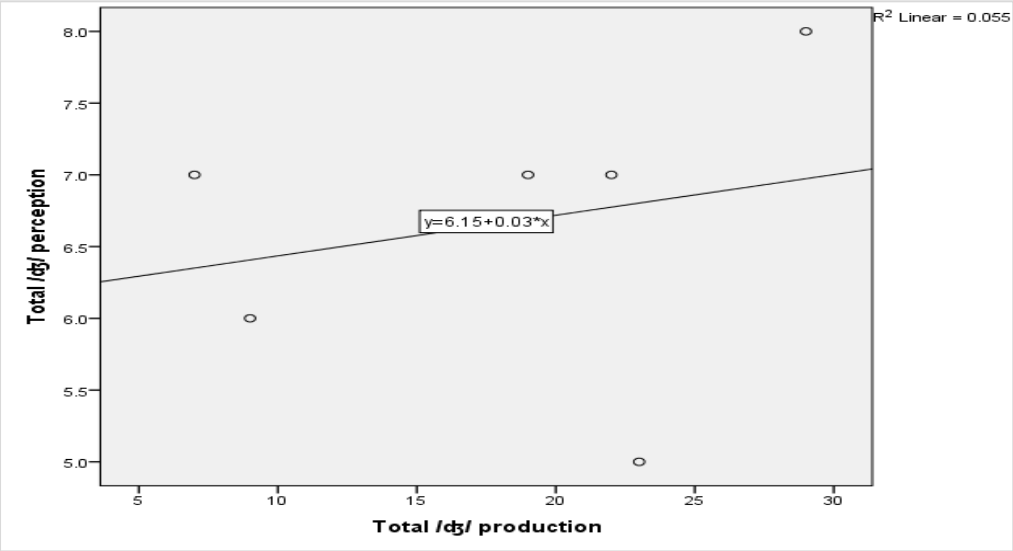
Table 10
The Correlation Between the Production and Perception of /ʤ/

		Total /ʤ/ perception
Total /ʤ/ production	Pearson Correlation	.234
	Sig. (2-tailed)	.655
	N	6

As demonstrated in Figure 6, the scatterplot is almost flat. This confirms that there is a weak correlation between the production and perception of the sound /ʤ/.

Figure 6

The Correlation Between the Production and Perception of /ʤ/



Moreover, the correlation between the production and perception of /ð/ was found to be insignificant and negatively weak, [r(-.088)= 4, p<.868] (see Table 11).

Table 11

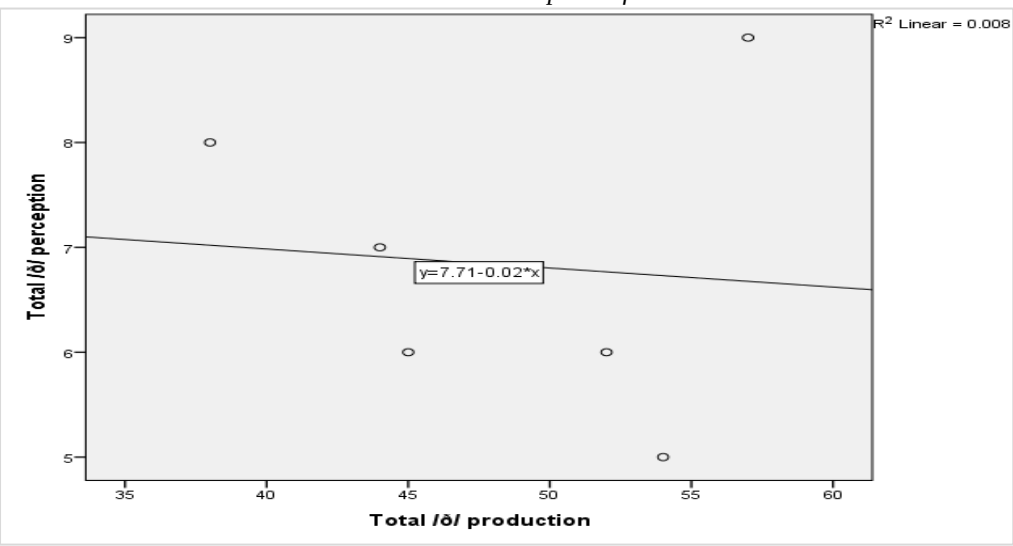
The Correlation Between the Production and Perception of /ð/

		Total /ð/ perception
Total /ð/ production	Pearson Correlation	-.088
	Sig. (2-tailed)	.868
	N	6

The nearly flat scatterplot, which can be observed in Figure 7, shows that there is a weak correlation between the production and the perception of /ð/.

Figure 7

The Correlation Between the Production and Perception of /ð/



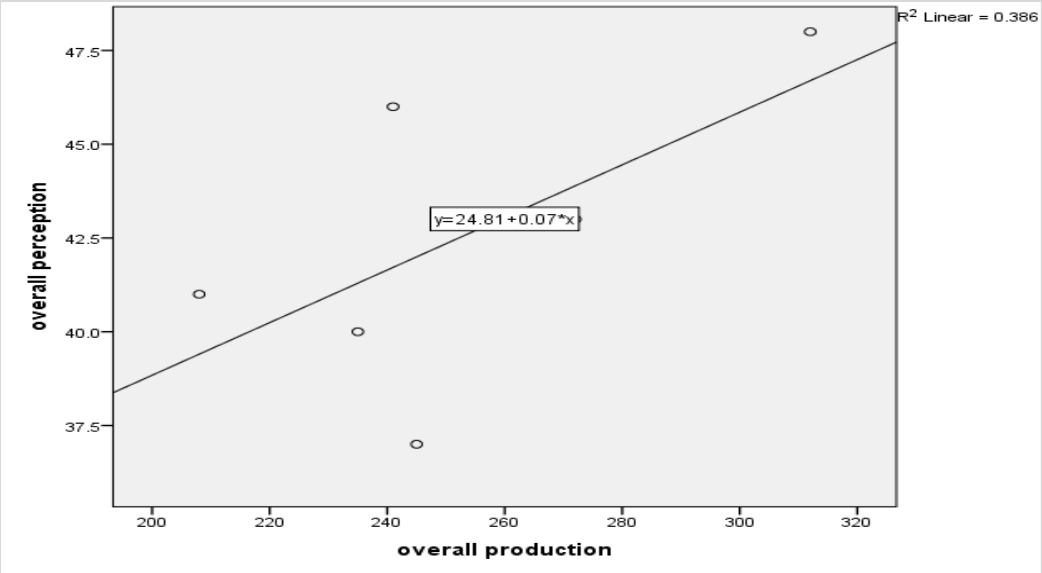
Additionally, an insignificant positive moderate correlation between the overall production and overall perception of the six target consonants was found, [r(.621)= 4, p < .188] (see Table 12).

Table 12
The Correlation Between Overall Production and Perception

		Overall Perception
Overall Production	Pearson Correlation	.621
	Sig. (2-tailed)	.188
	N	6

The scatterplot provides a summary of the results (see Figure 8). Generally, there was a positive moderate correlation between overall production and perception.

Figure 8
The Correlation Between Overall Production and Perception



DISCUSSION

The present study investigated the production and perception of six consonants of English language by Yemeni EFL learners. It was found that the easiest consonant sound produced and perceived by the six speakers was /θ/, while /ʤ/ was the most challenging sound for them to produce and perceive.

On the one hand, the participants had fewer problems perceiving the six target consonants. In just very few cases, the target consonants /p/, /v/, /θ/, /ð/, /ʧ/, and /ʤ/ were perceived as the Arabic sounds /b/, /f/, /t/, /d/, /ʃ/, and /ʒ/, respectively. In other words, as predicted by the PAM, the contrasted sounds were perceived as a single category assimilation only in very few cases (i.e., */p/-/b/* like /b/; */v/-/f/* as /f/; */θ/-/t/* as /t/; */ð/-/d/* like /d/; */ʧ/-/ʃ/* as /ʃ/; and */ʤ/-/ʒ/* like /ʒ/). These findings also supported Best and Tyler's (2007) PAM-L2 assumption that modifying existing phonological representations (e.g., /ʤ/) is more complicated than the acquisition of an L2 sound that has no direct counterpart in L1 (e.g., /tʃ/).

In the present study, the accuracy of perceiving the six target consonants is as follows: /θ/(89%)→/v/(80%)→/ʧ/(78%)→/p/and/ð/(76%)→ /ʤ/(74%).

In this study, the most troublesome sound perceived by Yemeni listeners was /ʤ/. Likewise, Evans and Alshangiti's (2018) study revealed that Saudi listeners with high and low proficiency faced the most difficulty in the identification of /ʒ/ followed by /ʤ/. The two groups identified the voiced affricate as /ɡ/ or /ʤ/. The difficulty in identifying the affricate /ʤ/ was unexpected since this sound is available in Arabic (Evans & Alshangiti, 2018). However, in Shafiro et al.'s (2012) study, the group with native Arabic found it challenging to identify /ð/ (65.8%) followed by /ʤ/(72.7%). Such low accuracy could be attributed to orthographic confusion between "j" (/ʤ/) as well as "ɡ" (/ɡ/) (Shafiro et al., 2012).

The current study also evidenced that Yemeni listeners encountered some difficulties in the perception of /ð/ and /p/, which were misperceived as /d/ and /b/, respectively, with a percentage of 24%. However, Saudi listeners with high proficiency had less difficulty identifying these two sounds, as Evans and Alshangiti (2018) revealed. Unlike this study, in some past studies, /ð/ was shown to be incorrectly identified as /v/ (e.g., Evans & Alshangiti, 2018; Shafiro et al., 2012) or /θ/ (e.g., Evans & Alshangiti, 2018). As for /p/, Evans and Alshangiti (2018) showed that low-proficiency participants' challenges with voicing perception were demonstrated in their achievement with the /p/-/b/ contrast since these listeners' performance with /p/ was good (74%), yet

/b/ was relatively poorly identified (68%) and primarily misidentified with as /p/ (27%). High-proficiency listeners experienced fewer difficulties with voicing, especially for the /p-b/ contrast, wherein identification was high (86% and 91%, correspondingly), demonstrating that with experience, learning to perceive this contrast is possible for Arab learners (Evans &Alshangiti, 2018). However, Shafiro et al. (2012) revealed that the Arab native participants produced /p/ accurately 99.8%, contradicting arguments in the literature on Arab learners' confusion of /p-b/ when producing /p/.

Furthermore, the findings of this study indicated that Yemeni listeners had less difficulty identifying /tʃ/ (see also Shafiro et al., 2012). In contrast, Evans and Alshangiti (2018) demonstrated low accuracy for /tʃ/ for low-proficiency listeners, and for this group, the most frequent errors were with /ʃ/ (29%). Yemeni listeners also had little trouble identifying /v/ compared to other sounds. Similarly, Evans and Alshangiti (2018) found that Saudi listeners with high and low proficiency identified /v/ accurately 91% and 87%, respectively. In Shafiro et al.'s (2012) study, native Arab listeners also perceived /v/ correctly with high accuracy (96.4%).

Besides that, the current study revealed that the least number of errors were detected in the identification of /θ/, which was misperceived as /t/ only 11% by Yemeni learners. However, Evans and Alshangiti (2018) revealed that /θ/ was correctly identified by high-proficiency and low-proficiency listeners, 74% and 54%, respectively. They found that Saudi listeners with high and low proficiency were shown to misidentify /θ/ mostly as /f/, yet also like /ð/; however, in a few cases, /θ/ was misidentified as /ʃ/, /v/, or /t/ by the low-proficiency group.

On the other hand, the participants of the current research struggled to produce the target consonants, particularly /ʤ/, followed by /p/, and then /v/. The accuracy of producing the target six consonants as descendingly ordered was as follows: /θ/(79%)→/ʧ/(73%)→/ð/(67%)→/v/(56%)→/p/(51%)→/ʤ/(25%). Numerous substitutions were discovered in producing the examined consonants by Yemeni EFL learners, primarily due to L1 transfer, as postulated by Gass and Selinker's (1992) Language Transfer Theory. English, as a foreign language, frequently interferes with Yemenis' native language, i.e., Arabic (Al-Hamzi et al., 2021; Bin Hadjah & Jupri, 2018). For instance, Arabic does not have /p/ (Al Abdey & Abdul-Rahman, 2021; Alwazna, 2020; Farrah & Halahlah, 2020; Siddig, 2022; Zoghbor, 2018). As a result, in some past studies, /p/ was produced as /b/ by Yemeni (e.g., Bin Hadjah & Hamzah, 2022; Hamzah & Bin Hadjah, n.d.) and other Arab learners of English (e.g., Farrah & Halahlah, 2020; Hamzah, Madbouly, et al., 2020). Several participants also wrongly produced /f/ instead of /v/ because Arabic lacks /v/ (Al Abdey & Abdul-Rahman, 2021; Alwazna, 2020; Farrah & Halahlah, 2020; Siddig, 2022; Zoghbor, 2018). Such a result was revealed by some researchers (e.g., Ababneh, 2018; Al Mafalees, 2020; Bin Hadjah & Hamzah, 2022; Farrah & Halahlah, 2020; Hamzah & Bin Hadjah, n.d.). Rehman et al. (2020) assured that Arab speakers frequently struggle with voicing contrasts (e.g., /v/-/f/; /p/-/b/).

With respect to the two fricatives of English, /θ/ was replaced with /t/ or produced with some voicing in other tokens. Such an incorrect production of /t/ for /θ/ was revealed by some past researchers in the production of Yemeni (e.g., Bin Hadjah & Hamzah, 2022; Mahfouz, 2013) and Arab learners of English (e.g., Khayra, 2017; Shalabi, 2017). However, a few Yemeni speakers substituted /ð/ with /θ/ or /d/. Such a result was also revealed by Mahfouz (2013) as well as Bin Hadjah and Hamzah (2022). Other Arab L2 learners of English were also evidenced to substitute /ð/ either with /θ/ (Farrah & Halahlah, 2020) or wrongly produce it as /d/ (Shalabi, 2017). It was also confirmed by Avery and Ehrlich (1992) that speakers of some dialects of Arabic might wrongly produce /d/ for /ð/ or /θ/ as /t/. Speakers of other dialects, however, might substitute /ð/ and /θ/ with /z/ and /s/, respectively. Nevertheless, /θ/ and /ð/ are typically less troublesome for Arab learners to produce, which could be due to their presence in their L1 (Zoghbor, 2018).

Additionally, some Yemeni speakers substituted /ʧ/ with /ʃ/, and such a finding was shown in other studies by Arab (e.g., Farrah & Halahlah, 2020; Hamzah & El-Weshahi, 2018) or Yemeni learners (Bin Hadjah & Hamzah, 2022; Bin Hadjah & Jupri, 2018; Hamzah, Bin Hadjah, et al., 2020). A few speakers wrongly produced the affricate /ʧ/ with some voicing. This could be due to mispronouncing the /t/ sound of /ʧ/ like the (ط) sound of Arabic, which might have been voiced, though this sound is currently articulated like a voiceless pharyngealised dental stop in the majority of dialects (Watson, 2002).

The most challenging sound produced by Yemeni speakers was /ʤ/, despite its availability in Arabic (Alwazna, 2020; Zoghbor, 2018). Several substitutions were detected in Yemeni speakers' production of /ʤ/, such as producing it inaccurately as /ʒ/, /ɡ/, /j/, /ʃ/, or producing it with no voicing. This may be attributable to the fact that various realisations are found in Yemeni dialects for the Arabic /ʤ/ (ج), including /ʤ/, /ɡ/, /j/, and /ɟ/ (Watson, 2002).

Yemeni EFL learners' substitution of /ʤ/ with /ʒ/ or /j/ was found by some previous researchers (e.g., Bin Hadjah & Hamzah, 2022; Hamzah, Bin Hadjah, et al., 2020), while Arab learners' replacement of /ʤ/ with /ɡ/ by was discovered in some prior studies (e.g., Tajeldeen, 2019; Thakur, 2020).

In addition, the current study revealed a significant effect of word position on the production

of the target sounds. Thus, H_1 was accepted. Such a result was also indicated by Alqarni (2013). However, word position did not significantly affect the perception of the target consonants. For this reason, H_2 was rejected. This result was congruent with Kelly and Keshishian (2019), who found no significant effect of word position on the perception of English sounds.

Moreover, a moderate correlation was found between overall production and perception. Therefore, H_3 was accepted. Such a moderate correlation between the production and perception of English sounds was indicated by some previous researchers (e.g., Berti et al., 2020; Hattori, 2010).

The present study is highly theoretical that may not be directly linked to learning, although results may have indirect implications for English language learning/teaching. According to the results of this investigation, it can be speculated that students may have problems producing and/or perceiving English sounds in certain word positions. Therefore, teachers may put a greater emphasis on specific sound environments. As shown by the results of this study, /p/, /v/, /ð/, and /ʤ/ were more challenging to produce when occurring word-finally as opposed to initial- and medial-word positions, whereas /θ/ and /ʧ/ were more difficult to produce in the initial word position in comparison to the other positions of a word, resulting in a difference in the occurrence of substitutions of the investigated sounds across the three-word positions (as detailed in Appendix A). The findings also demonstrated that the perception of the target sounds was slightly affected by word position. The sounds /p/ and /ð/ were more challenging to perceive in word-final position, whereas /v/, /θ/, and /ʤ/ were more difficult to perceive in word-initial position. The perception of /ʧ/ in the medial position was revealed to be less problematic than in word-final and word-initial positions.

The value of the present research lies in identifying specific word positions that create difficulties for L2 English learners when producing and perceiving certain English consonants. This may allow L2 learners of English to be more aware of such challenges when producing and perceiving English sounds in various sound environments. Moreover, several teachers should adopt a perspective on pronunciation as a vital and essential component of communication that ought to be integrated into classroom activities to improve EFL students' English articulation abilities (Almuslimi, 2020). Hence, this analysis may also assist EFL teachers in becoming more conscious of the value of Praat usage for pronunciation instruction since Praat software can be used to acoustically measure the features of segments (Boersma & Weenink, 2014). Furthermore, the present study may assist textbook designers in developing appropriate textbooks for teaching English to Yemeni EFL learners, especially that teaching pronunciation is somehow ignored in textbooks for teaching the English language in Yemen (Al-Hamzi et al., 2021).

This is the first investigation of its type to analyse how EFL Yemeni postgraduate students (at UUM) produce and perceive English sounds. The results have important implications for second-language learning as well as pronunciation instruction. Interestingly, the study results addressed a gap in the phonetic literature with regard to research on the relationship between the production and perception of English consonants by adult Arab L2 learners of English in general (and Yemeni EFL learners in specific), in addition to the effect of word position on production and perception.

CONCLUSION

The present study aimed to investigate how English consonants are produced and perceived by Yemeni EFL learners. More specifically, it examined the effect of word position on Yemeni EFL learners' perception and production of six English consonants, as well as the link between the two modalities. The production, but not the perception of the target sounds, was found to be significantly affected by word position. Furthermore, the Pearson correlation test showed an insignificant yet moderate positive correlation between the overall production and perception of the investigated consonants. Moreover, the present investigation provided a unique insight into Yemeni learners' production and perception of English consonants, shedding some light on specific challenging sound environments. The findings of this study can pave the way for future research into the production and perception of other English sounds, as well as the correlation between the two modalities, by testing the effect of other linguistic/non-linguistic factors on the perception and production of the investigated sounds.

REFERENCES

Ababneh, I. (2018). English pronunciation errors made by Saudi students. *European Scientific Journal, 14*(2), 244-261.

Al Abdey, A. A. W., & Abdul-Rahman, S. M. (2021). The effect of Iraqi EFL learners' proficiency level on their pronunciation of non-Arabic consonants. *Journal of Language and Communication*, *8*(2), 301-322.

Al-Hamzi, A., Al-Shrgabi, A., Al-Haidari, A., Faraj, M., & Al-Housali, T. (2021). Pronunciation errors of English front vowels by Yemeni EFL learners. *PAROLE: Journal of Linguistics and Education, 11*(1), 41–56.

Al Mafalees, F. (2020). Mispronunciation of English consonant sounds by Yemeni EFL learners at secondary schools: An analysis of the

problems and some remedies. *Language in India*, *20*(1), 108–121.

Almuslimi, F. (2020). Pronunciation errors committed by EFL learners in the English department in faculty of education-Sana'a university. *Millennium Journal of English Literature, Linguistics and Translation*, *2*(1), 51-72. https://doi.org/10.47340/mjellt.v1i2.5.2020

Alqarni, A. (2013). *The realization for the English voiceless postalveolar affricate /tʃ/ in Najdi Saudi ESL learners production* [Master's thesis, Southern Illinois University].

Alwazna, R. (2020). The differences in consonantal pronunciation between formal English and Saudi Hijazi English and their implications for oral intralingual translation. *Journal of psycholinguistic research*, *49*(4), 571-582.

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

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation.* Oxford University Press.

Baagbah, S., Jaganathan, P., & Mohamad, A. (2016). Investigating the challenges of acquiring/ð/and/v/English sounds by Yemeni EFL learners. *Imperial Journal of Interdisciplinary Research*, *2*(7), 678-682.

Baese-Berk, M. M. (2019). Interactions between speech perception and production during learning of novel phonemic categories. *Attention, Perception, & Psychophysics*, *81*(4), 981-1005.

Berti, L. C., Guilherme, J., Esperandino, C., & de Oliveira, A. M. (2020). Relationship between speech production and perception in children with speech sound disorders. *Journal of Portuguese Linguistics*, *19*(1), 1-13. https://doi.org/10.5334/jpl.244

Best, C. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception: New directions in research and theory. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171–204). York Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775-794.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 13–34). John Benjamins.

Binasfour, H. (2018). *Investigating the perception and production of the Arabic pharyngealised sounds by L2 learners of Arabic* [Doctoral dissertation, University of Reading].

Bin Hadjah, N. A. K., & Hamzah, M. H. (2022). Phonological processes in the production of English consonants by Yemeni EFL speakers of English. *Res Militaris*, *12*(2), 334-355.

Bin Hadjah, N. A. K., & Jupri, R. (2018). Pronunciation of the English voiceless postalveolar affricate /ʧ/ in Yemeni EFL learners' production. *ASIAN TEFL*, *3*(2), 103–117.

Black, K. (2019). *Sports and physical education.* ED-TECH PRESS.

Boersma, P., & Weenink, D. (2014, June 21). *Praat: Doing phonetics by computer*. https://www.researchgate.net/publication/259810776

Boersma, P., & Weenink, D. (2022, March 17). Praat (version 6.2. 10) [Computer software]. http://www.fon.hum.uva.nl/praat

Chan, A. (2011). The perception of English speech sounds by Cantonese ESL learners in Hong Kong. *TESOL Quarterly*, *45*(4), 718–748.

Chao, S. C., Ochoa, D., & Daliri, A. (2019). Production variability and categorical perception of vowels are strongly linked. *Frontiers in Human Neuroscience*, *13*, 1-9. https://doi.org/10.3389/fnhum.2019.00096

Cheng, H. S., Niziolek, C. A., Buchwald, A., & McAllister, T. (2021). Examining the relationship between speech perception, production distinctness, and production variability. *Frontiers in Human Neuroscience*, *15*,1-12. https://doi.org/10.3389/fnhum.2021.660948

Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative Research*. Pearson.

Culleton, T. (2021, April 16). *Production and perception of English vowels by second language learners*. Retrieved March 11, 2021, from https://dc.uwm.edu/uwsurca/2021/asynchronous/29/

Dancey, C., & Reidy, J. (2007). *Statistics without maths for psychology*. Pearson education.

Docio-Fernandez, L., & Garcia-Mateo, C. (2015). Speech production. In S.Z. Li, & A.K. Jain

(Eds.), *Encyclopedia of Biometrics*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7488-4_199

Emran, A., & Anggani, D. (2017). The errors of segmental phonemes among Libyans English students studying in Semarang city, Indonesia. *Language Circle: Journal of Language and Literature*, *11*(2), 183–197.

Evans, B., & Alshangiti, W. (2018). The perception and production of British English vowels and consonants by Arabic learners of English. *Journal of Phonetics*, *68*,15–31.

Farrah, M., & Halahlah, N. (2020). Pronunciation problems among Palestinian English major students in Hebron University. *International Journal of Arabic-English Studies, 20*(1), 203–226.

Firdaus, S., Indrayani, L., & Soemantri, Y. (2020). The production of interdental fricatives by English as a foreign language students in English Course Bandung. *Linguistics and ELT Journal*, *8*(1), 1–9.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In S. Winifred (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3-83). Cambridge University Press. https://doi.org/10.1017/9781108886901.002.

Gass, S., & Selinker, L. (1992). *Language transfer in language learning*. John Benjamins.

Hagiwara, R. (2009, November 18). *How to read a spectrogram*. Retrieved October 23, 2020, from https://home.cc.umanitoba.ca/~robh/howto.html

Hamzah, M. H. (2013). *The acoustics and perception of the word-initial singleton/geminate contrast in Kelantan Malay* [Doctoral dissertation, The University of Melbourne].

Hamzah, M. H., & Bin Hadjah, N. A. K. (n.d.). *Production of English consonants by Yemeni EFL learners of English: The case of /p/ and /v/* [Manuscript submitted for publication].

Hamzah, M. H., Bin Hadjah, N. A. K., & Abdullah, A. H. (2020). The production of English affricates by Yemeni EFL learners of English. *The Asian Conference on Language 2020, Tokyo, Japan*, 1–16.

Hamzah, M. H., & El-Weshahi, O. (2018). Deaffrication process among Arab learners of English: The case of voiceless postalveolar affricates /ʧ/. *11th Language for Specific Purposes International Conference & 10th Global Advances in Business Communication Conference (LSP-GABC 2018)*.

Hamzah, M. H., Madbouly, A. E. S., Halim, H. A., & Abdullah, A. H. (2020). The production of the English stop voicing contrast by Arab L2 speakers of English. *Indonesian Journal of Applied Linguistics*, *10*(2), 434-444. https://doi.org/10.17509/ijal.v10i2.28615

Hattori, K. (2010). *Perception and production of English/r/-/l/by adult Japanese speakers* [Doctoral dissertation, University College London].

Huensch, A. (2013). *The perception and production of palatal codas by Korean L2 learners of English* [Doctoral dissertation, University of Illinois at Urbana-Champaign].

Huwari, I. (2019). Pronunciation errors in Egyptians' English. *Literary Endeavour*, *10*(5), 31-44.

Jevring, C. (2015). *I perceive, therefore I produce?: A study on the perception and production of three English consonantal sounds by Swedish L2 learners* [Master's thesis, Stockholm University].

Junker, F. B., Schlaffke, L., Bellebaum, C., Ghio, M., Brühl, S., Axmacher, N., & Schmidt-Wilcke, T. (2020). Transition from sublexical to lexico-semantic stimulus processing. *Frontiers in Systems Neuroscience*, *14*, 1-8.

Kadiri, G. C., Adeyi, V., Ekwueme, J., & Samaila, Z. B. (2020). The production and perception of the affricate/tʃ/and the fricative/ʃ/by Igala ELS users. *Theory and Practice in Language Studies*, *10*(1), 19-25. http://doi.org/10.17507/tpls.1001.03

Kaewchum, C. (2018). *A study on Thai kindergarten teachers' perception and production of 10 English problematic final consonant sounds: A case study in the Northeast of Thailand* [Master's thesis, Thammasat University].

Kelly, N. (2019). The perception of dental and alveolar stops among speakers of Irish English and American English. *English Language & Linguistics*, *23*(2), 277–302. https://doi.org/10.1017/S1360674317000405

Kelly, N., & Keshishian, L. (2019). The voicing contrast in stops and affricates in the Western Armenian of Lebanon. *INTERSPEECH*, 1721–1725.

Khayra, D. (2017). *The influence of the mother tongue (Arabic) on the pronunciation of the students of English* [Master's thesis, University Abdelhamid Ibn Badis].

Kochetov, A. (2004). Perception of place and secondary articulation contrasts in different syllable positions: Language-particular and language-independent asymmetries. *Language and Speech*, *47*(4), 351–382.

https://doi.org/10.1177/0023830904047004020 1

Lee, Y. K. (2019). Korean college learners' sound perception and production of English voiceless fricative. *언어과학연구*, 89, 199–219. https://doi.org/10.21296/jls.2019.6.89.19 9

Lengeris, A., & Nicolaidis, K. (2016). The identification and production of English consonants by Greek speakers. *Selected Papers on Theoretical and Applied Linguistics*, *21*, 224-238. https://doi.org/10.26262/istal.v21i0.5227

Lersveen, L. R. (2018). *The perception and production of nonnative English consonants in native Norwegian speakers* [Master's thesis, NTNU].

Mahfouz, T. (2013). *Yemeni dialect*. Lulu Enterprises Incorporated.

Maiunguwa, A. (2015). *Perception and production of English fricatives by Hausa speakers* [Master's thesis, University of Malaya].

Nunnally, J. (1978). *Psychometric theory* (2nd ed.). McGraw Hill.

Nurfitriani, E. (2019). *The difficulties of producing English consonant sound for EFL Sundanese students (A case of the eleventh grade students of SMA Negeri 1 Ciamis in the academic year 2018/2019)* [Undergraduate thesis, Universitas Negeri Semarang].

Oh, E. (2019). Effects of gender on the use of voice onset time and fundamental frequency cues in perception and production of English stops. *Linguistic Research*, *36*(1), 67–89. https://doi.org/10.17250/khisli.36.1.201903.00 3

Pallant, J. (2020). *SPSS Survival Manual: A step by step guide to data analysis using IBM SPSS*. Taylor & Francis.

Pei, Y. (2022). Perception and production of English consonants by senior high school students. *Journal of Education and Development*, *6*(3), 47- 52.

Rehman, I., Silpachai, A., Levis, J., Zhao, G., & Gutierrez-Osuna, R. (2020). The English pronunciation of Arabic speakers: A data-driven approach to segmental error identification. *Language Teaching Research*, 1–27. https://doi.org/10.1177/136216882093188

Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics*, *39*(1), 187-224.

Schmitz, J., Díaz, B., Fernandez Rubio, K., & Sebastian-Galles, N. (2018). Exploring the relationship between speech perception and production across phonological processes, language familiarity, and sensory modalities. *Language, Cognition and Neuroscience*, *33*(5), 527-546.

Seo, M., & Lim, J. (2016). Korean EFL learners' production and perception of English sound contrasts. *외국어교육*, *23*(1), 111–132

Shafiro, V., Levy, E. S., Khamis-Dakwar, R., & Kharkhurin, A. (2012). Perceptual confusions of American-English vowels and consonants by native Arabic bilinguals. *Language and speech*, *56*(2), 145-161. https://doi.org/10.1177/0023830912442925

Shalabi, M. (2017). Think, tink or sink, the phonological awareness of English voiceless interdental fricative [θ] and [ð] among Chinese, Arab and Pakistani learners of English. *American Journal of Education and Information Technology*, *1*(3), 31–37.

Siddig, O. A. (2022). Contrastive analysis of most relevant features of English and Arabic languages. *Journal of Positive School Psychology*, *6*(7), 4198-4210.

Sioson, I. C., & Chang, K. R. (2017). The relationship between perception and production of fricatives by Thai learners of English. *6th International Conference on Language, Innovation, Culture and Education 18th& 19th February, 2017*, 27–41.

Stasenko, A., Bonn, C., Teghipco, A., Garcea, F., Sweet, C., Dombovy, M., McDonough, J., & Mahon, B. (2015). A causal test of the motor theory of speech perception: A case of impaired speech production and spared speech perception. *Cognitive Neuropsychology*, *32*(2), 38–57. https://doi.org/10.1080/02643294.2015.103570 2

Sulistyorini, D., & Wibowo, R. (2021). An analysis of students perception and production problems of pronouncing English palato alveolar sounds. *Marine Science and Technology Journal*, *1*(2), 96–111. https://doi.org/10.31331/maristec.v1i2

Syed, N. (2011). Perception & production of consonants of English by Pashto speakers. *Journal of Humanities & Social Sciences*, *19*(1), 119–146.

Syed, N. A., Ansari, S., & Gopang, I. B. (2017). Perception and production of consonants of English by Pakistani speakers. *International Journal of English Linguistics*, *7*(3), 201-214. http://doi.org/10.5539/ijel.v7n3p201

Tajeldeen, H. A. I. (2019). *The influence of pronunciation errors in changing the meaning of words among EFL learners* [Master's thesis, Sudan University of Science and Technology].

Thakur, V. S. (2020). Phonological problems of Omani EFL learners: Pedagogical perspectives and implications. *Arab World English Journal*, *11*(1), 29–43.

Watson, J. C. E. (2002). *The phonology and morphology of Arabic*. Oxford University Press.

Zhang, A., Feng, H., Wang, S., & Dang, J. (2016, October). Relationship between perception and production of English vowels by Chinese English learners. In *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)* (pp. 1-5). IEEE. https://doi.org/10.1109/ISCSLP.2016.7918479

Zhang, B., Zhang, J., & Lee, S. H. (2021). Perception and production of English fricatives by Chinese learners of English: Error patterns and perception-production relationship. *Phonetics and Speech Sciences*, *13*(1), 25-36.

Zoghbor, W. S. (2018). Teaching English pronunciation to multi-dialect first language learners: The revival of the Lingua Franca Core (LFC). *System, 78*, 1-14.

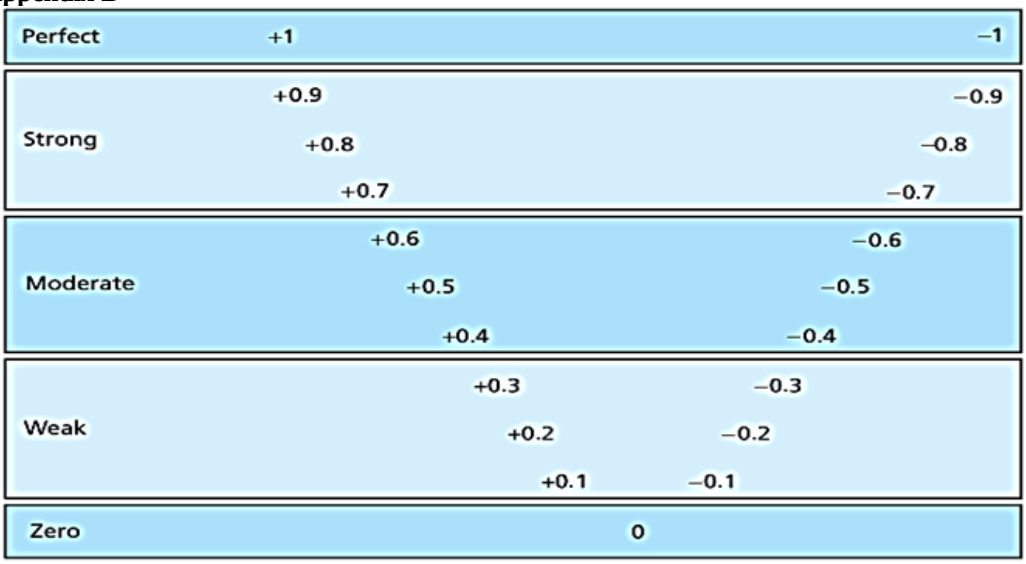
APPENDICES

Appendix A

Substitutions of the Target Sounds by Each Speaker Across the Three-Word Positions

Speakers	Word Position	Target Sounds					
		/p/	**/v/**	**/θ/**	**/ð/**	**/ʧ/**	**/ʤ/**
MS1	**I**	/b/	/f/	_____	/d/	/ʃ/	with no voicing
	M	/b/	/f/	/t/, with some voicing	with no voicing	/ʃ/	with no voicing, /ʒ/, /g/, /j/
	F	/b/	/f/	/t/	/d/, /t/, /θ/, with no voicing	/ʃ/, with some voicing	with no voicing, /ʒ/
MS2	**I**	/b/	/f/	_____	/θ/, with no voicing	_____	/ʒ/, /ʧ/, with no voicing
	M	_____	/f/	_____	with no voicing	/ʃ/	/ʒ/, /k/, /g/, with no voicing
	F	/b/	/f/	_____	/θ/, with no voicing	/ts/, with some voicing	with no voicing
MS3	**I**	/b/	/f/	/t/, with some voicing	/d/, /t/	/ʃ/	/ʒ/, /g/
	M	/b/	_____	/t/, with some voicing	/t/	/ʃ/	/ʒ/, with no voicing
	F	/b/	/f/	/t/, deleted	/t/, /θ/, deleted	/ʃ/, deleted	/ʒ/, with no voicing
FS4	**I**	/b/	_____	with some voicing	/d/, /θ/	_____	with no voicing
	M	/b/	/f/	with some voicing	_____	_____	/ʒ/, with no voicing
	F	/b/	/f/	with some voicing	/θ/	with some voicing	with no voicing
FS5	**I**	/b/	/f/	with some voicing	/θ/, deleted	/ʃ/	/g/, with no voicing
	M	/b/	/f/	/ð/	_____	/ʃ/	/j/, with no voicing
	F	/b/	/f/	_____	/θ/, with no voicing	/ʃ/	/g/, with no voicing
FS6	**I**	/b/	/f/	with some voicing	/θ/	_____	/ʒ/, /g/, with no voicing
	M	/b/	/f/	_____	with no voicing	_____	/j/, with no voicing
	F	/b/	/f/	/s/	/θ/, with no voicing	_____	with no voicing

Appendix B



The strength of negative and positive correlation coefficients
Source: (Dancey & Reidy, 2007)