Available online at www.jlls.org

JOURNAL OF LANGUAGE AND LINGUISTIC STUDIES

ISSN: 1305-578X

Journal of Language and Linguistic Studies, 18(Special Issue 2), 1194-1207; 2022

Voicing contrast of L2 final stops: A case study on ESL learners from Saudi

Abdullah Alotaibi [a 1]

[a] *Majmaah University, Majmaah, Saudi Arabia*

APA Citation:

Alotaibi, A. (2022). Voicing contrast of L2 final stops: A case study on ESL learners from Saudi. *Journal of Language and Linguistic Studies, 18*(Special Issue 2), 1194-1207.

Submission Date:17/08/2021

Acceptance Date:03/12/2021

Abstract

Previous research in L2 perception has shown limited and diverge findings for final stop voicing contrasts by learners with different L1 backgrounds in comparison to natives of the target language. This paper aimed to investigate L2 learners' perceptual identification of voicing in English final stops based on the duration of succeeding vowel, and the impact of prior linguistic expertise in establishing L2 phonetic category as learners' level of proficiency develop. The subjects were adult Arabic L2 learners (N = 9) with different proficiency levels (advance, intermediate, & beginners). The subjects' performance was later compared to controlled native English group (N=3). Data were elicited through listening to stimuli in one task that contained three controlled conditions and one test condition. The analysis employed descriptive statistics, and one-way ANOVA. The findings showed that as the subjects' proficiency level developed, their accuracy in identifying final stops voicing increase. It also showed the possibility of establishing L2 phonetic category that seemed developed in correspond to the development of L2 proficiency level. The results illustrate the systematic processing of L2 sounds which have implications for theories in L2 perception.

Keywords: L2 phonology; final stops; voicing; pseudowords; perception; phonological category; voicing contrasts; vowel duration

1. Introduction

Perception in the second language, for adult L2 learners, is important as production to reach intelligibility in L2. As L2 production could be influenced by learners' L1 (Alotaibi, 2018; Flege, 1980; Flege & Port, 1981; Flege & Eefting, 1987), the L2 perception could also be vulnerable to such influence (Darcy & Krüger, 2012) The Perception Assimilations Model (PAM) by (Best, 1995), proposes that L2 targeted sounds could be in one way or another assimilated by L1 sounds during the perception process. Similarly, the Native Language Magnet theory (NLM) by (Kahl, 2000), suggests in the perception process of L2, L1 may act as a filter in perceiving L2 sounds, which would change the property of the perceived sounds. However, it is worth noting here that such an influence could be sparse through language experience with L2 (Al Malwi, 2017). The impact of language experience in L2 speech has a resilient ground that was endorsed by speech learning model. The Speech Learning Model (SLM) (Flege, 1995) predicts that increase in language experience in L2 may lead to better perception for L2 sounds, which is due to forming new phonetic category for the new sound. Though if there was an equivalent classification to the targeted sounds, it may hinder the formation of the new

[1] Corresponding author.
E-mail address: an.otaibi@mu.edu.sa

sound category for sounds that share similar acoustics features between L1 and L2 (e.g. Alotaibi, 2013; Alotaibi, 2018).

In the light of these empirical studies and theories, the present paper aims examining the probability of forming L2 phonological category beside L1 phonologcal category and the impact of language experience in forming the new L2 category as learners' level of proficiency develops. In order to do so, this paper investigates the underlying voicing contrasts, through AX discrimination task, in the perception of English final stops (e.g., lab / lap, made /mate, sight / side) by Adults Arabic L2 English learners with different proficiency levels. Since there are certain phonological contrasts between subjects' L1 Arabic and L2 English (Al-Ani, 1970; Aljaser, Jackson, Vitevitch & Joan, 2018; Alotaibi, 2013), it is expected that there could be a variation emerge in the course of this experiment. Furthermore, this paper tended to examine the facility of Arabic adult L2 learners in correctly identifying stops voicing in word final according to the acoustic cue of the preceding vowel length. It is hoped that the existing research would provide a helpful illustration of the systematic processing by which voiced English stops are perceived by adult bilingual listeners.

1.1. Literature Review

1.1.1 Previous studies on non-Arabic ESL learners

In the last few decades, a substantial number of studies examined the phonological processing of English stops by L2 learners. However, a scarce number of these studies have considered the language experience, acoustics cues, and phonological environment in their investigation. One of these early studies was conducted by Flege & Efting (1987) to investigate the perception of English stop category through an identification task on ESL native speakers of Spanish. The study found that the overall performance was varying within-subjects according to their language experience and age of exposure to L2. Specifically, accuracy increased on those who encountered to L2 consistently and at an early age than older age. When Flege & Efting (1987) compared the Spanish subjects to similarly-aged English monolinguals, the subjects were found to have developed an L2 phonetic category for targeted sounds. Yet, it seemed to be influenced more by their native language phonetic category, which could be due to filtering the perceived L2 sounds through L1 phonological knowledge, as proposed by NLM (Kuhl, 2000). This may lead to construct the newly established phonetic category under the notion of interphonolgy. Major (1987) pinpointed on the issue of phonological interference that is readily evident in the initial stage of language acquisition. However, Major (1987) asserted that through what he called "developmental factors", a learner's acoustical skills would steadily develop toward a native-like phonological system, constrained by the degree of linguistic exposure. Therefore, the consistent experience with L2 for the subjects in (Flege & Efting, 1987) appears, to a greater extent, to promote the formation of L2 phonetic category to be near native-like.

In a similar vein, Flige, Munro, and Macky (1995) found language experience and age of onset in learning L2, for Italian ESL learners, effects on acquiring L2 English stops in a native-like manner. Likewise, along with language experience, age appeared here to play a role in L2 phonology, but to a lesser degree; adding to the common claims for the effect of age as a factor on attaining native-like phonological perception (Ohde & Abou-Khalil, 2001). Nevertheless, other studies disagree with the overriding determinant of age on L2 phonological processing (Buali, 2010; Hazan & Boulakia, 1993). A respective number of studies illustrate that L2 learners, through extensive experience in the target language, could enhance their phonological skills (Zampini & Green, 2001) especially in contrasting between voiced and voiceless stops (Mora, 2007). It is, therefore, a matter of input quantity and quality rather than the age of exposure, because early acquisition does not necessarily mean native-like perception with the absence of enough input (Trofimovich, 2008). In a recent study carried on German L2 adult English learners, Smith, Hayes, Bruss, and Harkir (2009) discovered that ESL learners with

high language experience succeeded in developing an L2 phonological category that helped them to contrast voicing on stops in word-final position, which comes in line with similar finding on Chines L2 English learners (Flege, McCutcheon & Smith 1987). In Smith et al., (2009) study, although the German phonological system manifests a pattern of stops devoicing in word-final position (Charles-Luce,1985), it was conceivable for the Subjects to effectively acquire near native-like voicing contrast in L2 English after adequate exposure to L2 English phonological system.

1.1.2 Previous studies on Arabic ESL learners

In the Arabic language, as in German, there is a similar tendency toward devoicing of stops in word-final position (Al-Ani, 1970; Watson, 2002). Accordingly, with considering Major (1987) viewpoint on L2 learners phonological development stages, one could postulate that since voiced obstruent are devoiced in word-final position in Arabic, Arabic ESL learners, specifically those in an early stage of acquisition with less interaction with L2, may encounter some challenges in contrasting voicing on stops in word-final position. Flege (1980), and similarly Flege and Port (1981) examined voicing contrast for stops in word-final by Saudi L2 English learners. The results showed that even though the influence of L1 phonological system was salient, the sufficient amount of exposure to L2 had helped the subjects to contrast targeted sounds with a high rate of accuracy, differently from subjects with lower amount of exposure to L2. Consistent exposure to L2 in both studies seems helped learners to develop a new phonemic category meant for L2 that contains target like phonetic features, which is absent from L1 and to some extend suppresses the influence of L1 phonological system. However, in both studies Flege (1980), and Flege and Port (1981), the discussion falls short to discuss thoroughly the acquisition of some of the phonological cues that could help in contrasting voicing in stops. In a later study by Fokes, Bond, and Steinberg (1985) on twelve Arabic children learning English as L2, the researchers looked at their ability to contrast of /p/ vs /b/ and /t/ vs /d/ in a syllable-coda position that was cued by vowel duration preceding these four targeted phonemes. It's worth mentioning here that in the English phonological system there is a 'temporal compensation' in vowels (Chen, 1970). That is vowel duration changes according to the following consonant; if it precedes voiced consonant it becomes longer and gets shorter before a voiceless consonant. Such accommodating in vowel length to neighbouring consonant is a strong cue to contrast voicing for the following stops in word-final (Flege, Hllengbrand, Ingrisano & Smith 1984; Lehman & Sharf, 1989; Ladefoged, 2006; Krause, 1982; Mayo & Turk, 2004). In Fokes et al., (1985) study, it has been reported that after some language experience, the subjects succeeded to perform near-native speaker-like, and had acquired one of L2 phonological specific features, which is the vowel duration cue, that helped them to contrast voicing in stops. What was most interesting most in Fokes et al., (1985) study, is even though Arabic has no voicing dependent distinction for stops (Flege, 1979), subjects were able to utilize the preceding vowel length in the perception process for L2 English stops. In Arabic, a vowel has a random phonological pattern that does not depend on the neighbouring consonant for its length (Port, Al-Ani & Maeda, 1980). Being able to acquire L2 specific phonetic feature, e.g. the preceding vowel duration, is quite interesting and illustrate the critical role of language experience on the possibility of establishing L2 phonetic category. Yet, the newly established phonetic category may seem not fully separate from L1 as was suggested by SLM (Flege, 1995). Therefore, it's important to remember that acquired L2 phonetic specific features may not fully resemble native speaker like. It is very well-documented that L2 English sounds, to certain degree within Arabic subjects, tend to be neutralized in subjects' perception or production in their underlying voicing contrast (i.e., made /mate [m>t / m>t]; sight / side [s>t / s>t]) in adult (Buali, 2010; Flege & Port, 1981) as well as in children (Fokes et al., 1985). This reflects the fact that in the previously conducted studies on Arabic L2 learners, there could be non-phonological factors that may exert an effect on the perception of targeted sounds, such as lexical factor, orthography, and word frequency. Therefore, taking these factors into

account in this investigation may help to have a crystallized understanding of the formation of L2 phonetic category and the subject's performance under a more controlled experiment. The finding of previous studies to some extent seemed consistent with the earlier prediction by SLM (Flege, 1995) regarding the establishment of L2 phonetic category. Which gives solid ground to hypothesize that proficiency level would have a strong reflection on overall subjects' performance, in which subjects with a high level of proficiency would outperform those with low level due to better establishment of L2 phonetic category. Apparently, more focused studies is needed to examine L2 perception and the ability of bilinguals to identify various sounds among their L1 and L2 phonemic space. With the observed influence of language experience in the literature, a question was raised on whether more controlled and intensive language experience in L2 could in the long-term limit the robust influence of L1's phonetic category on L2, and lead L2 learners to perform native speakers like in perceptual level. Also, it would be quite exciting to look at the utilizing of pre-consonant vowel acoustic length features, and the possible use as a hint to voicing by adult Arabic L2 learners. Similarly, as Fokes et al. (1985) did with children; but this time on adults through different learning stages in L2 English. It helps in determining the degree to which learners may have begun to develop L2 phonetic category in their underlying phonological system.

2. Method

This paper was organized in the frame of quantitative research, and the research questions were:

1- Do Adult L2 learners utilize vowel acoustical length variation as a hint in identifying voicing of final stop?

2- Do adult L2 Learners of English establish new phonetic category besides their L1? And does such establishment of L2 phonetic category develop as learners' level of proficiency increase?

2.1. Participants

The subjects were divided into 2 groups: the control group was American native English speakers, and the experimental group was Saudi Arabian ESL learners, Total N= (12). The control group subjects consisted of three undergrads (1 female and 2 male) with an age range between 19 and 25 and a Mean age of 22. The experimental group consisted of 9 subjects (4 females and 5 males), with age vary between 19 and 23 and a Mean age of 21. The experimental group subjects were chosen among ESL learners who were studying in United State and had stayed in the US from 3 to 24 months. Based on the participants' self-reported TOEFL IBT scores, their proficiency levels fell within the range of beginner, intermediate, and Advance learners. For more detail in demographic information see Table 1 and for proficiency level criteria see Table 2.

Table 1. Mean of Subjects' demographic information

	Native	Advance	Intermediate	Beginners
Mean age	22	21	21	21
Mean length of residence -	24 month	18 month	6 month	
Mean L1 language use	-	% 30	% 50	% 70
Mean L2 language use	-	% 70	% 50	% 30
Mean L2 learning home	-	12 years	12 years	12 years
Female	2	1	1	2
Male	1	2	2	1

Table 2. Proficiency level criteria

Level of proficiency	IBT-TOEFL Score
Advance	TOEFL 80-120 / Level 6 & 7
Intermediate	TOEFL 50-70 / Level 3, 4 & 5
Beginners	TOEFL 20-40 / Level 1 & 2

2.2. Instrument(s)

It contained the target phonemes preceded by vowel with different durations long / short (see Figure 1, 2). It aimed to examine the subjects' ability in using vowel length as a cue to distinguish the differences between the target phonemes and their counterparts' ones that are produced in the same place of articulation. Similar to condition 1, the vowel length was set as follow: short vowel length were 100 (ms), and 300 (ms) for long vowels. The stimuli was designed as follow:

The subjects heard the words with final voiced stop, but the preceding vowel duration varied as follows:

Short vowel+ Voiced stop

Long Vowel+ Voiced stop

2.3 Procedure and Analysis

Since this study focused on the perception of the subjects, each one of the subjects was examined individually. The total time took 25 minutes for each subject to finish the whole experiment. Prior to the start, the subjects had been given detailed instruction and went through a brief familiarizing task near-like the actual tasks. Then the subjects were asked to fill up a questionnaire form. The Stimuli was divided mainly into two parts: demographic information and perception task. The demographic information was collected through a questionnaire which elicited the information that included: (a) educational status; (b) age; (c) gender; (d) length of learning English in homeland; (e) length of learning English in the USA; (f) other spoken languages beside Native language and English; (g) length of living in the USA.

The perception task included listening to pseudowords that have a natural male voice and were recorded on high quality for over headphones computer-based task at a comfortable level. On the screen, there was a running software named Praat (Boeirsma & Wenink, 2017) that was customized to display two response options consisting of two words beside the played audio. The subjects were given a computer keyboard that has two specified function keys to click on and choose only one of them as an answer choice. These function keys correspond to the displayed words on the computer screen. The subject can click on the function keys right after they hear the played audio. The subject heard disyllable English pseudowords through four blocks. Three of the four blocks were controlled conditioned; one for vowel length identification task, the second for phonemic identification task in word-initial position, and the third was also for phonemic identification task but in word middle position. The fourth block was the test condition and it contained the target phonemes in pseudowords, in the final position in the word. Each block was presented at a time through the computer screen (24 V length + 12 word-initial +12 word middle + 12 word final). In each block, the trials of pseudowords were counterbalanced, and the audio was randomized. The subjects had to go through an AX discrimination task for the targeted sound in each block with four possible trial orders (AB, BA, AA, BB). The interstimulus interval ISI was 600ms and the trials were played with 1000ms waiting after the response.

2.3.1 Controlled condition 1

The first condition was designed to examine the subjects' capability to identify the variations between long and short vowels (see Table 3). Having such vowel length distinction by the subjects would let the examiner know that the identification of vowel length is acquired by the subject and therefore consider the possibility of using it as a prompt to the voicing of the targeted consonant. However, failing in making distinguishing, would lead the examiner to expect that the subject may disregard the vowel length as a clue to the voicing, and built their decision based on their intuition. Since long vowels are two third longer than short vowels (Mack, 1982), the short vowel length was set at 100 (ms), and 300 (ms) for long vowels (see Figure 1, 2). The vowels in the words were shortened by manipulating the original vowel length through Praat. The native should be able to tell the difference between the long and short vowel, if they could not that means that the stimuli is invalid.

Table 3. Pseudo words for vowel length identification

[ki:n]	[bun]	[pæm]	[ti:n]	[dim]	[gu:m]
[bu:n]	[kin]	[di:m]	[gum]	[tin]	[pæ:m]

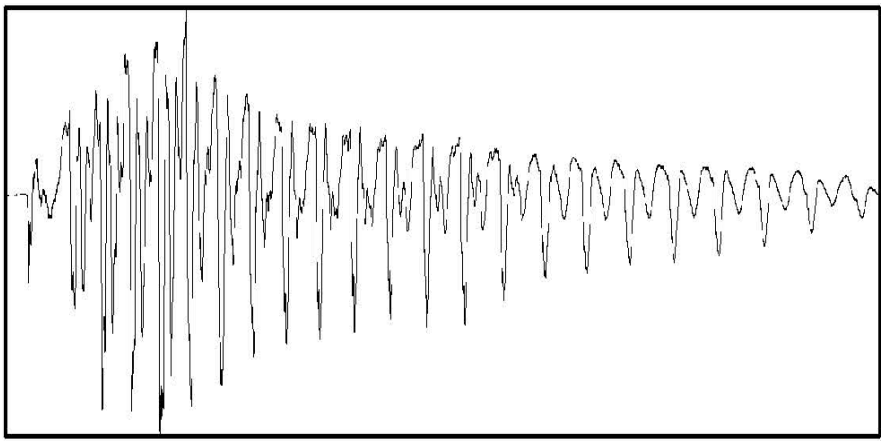

Figure 1. Spectrogram for [dim] with 100 ms length.

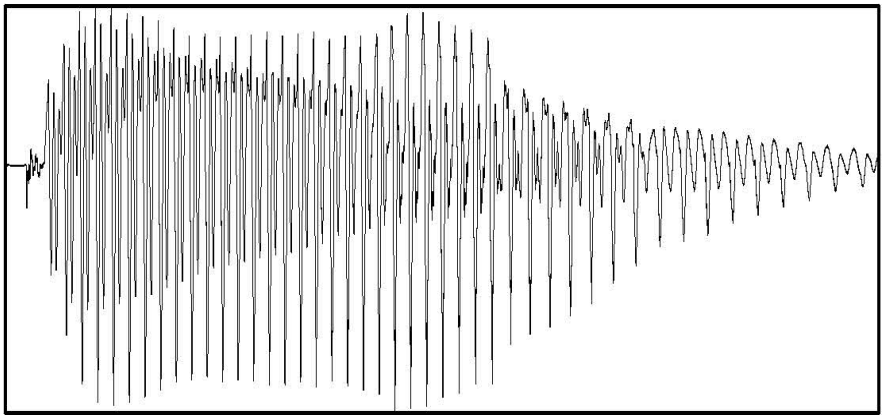

Figure 2. Spectrogram for [dim] with 300 ms length.

2.3.2 Controlled condition 2 & 3

The second condition is designed to examine the subjects' ability to identify the distinctions between voiced and non-voiced stops that are articulated in the same articulation point. The vowel length here is not important, and the aim of having this condition is to see if the subjects have acquired such distinction between these phonemes or they just end up perceiving the two phonemes that are produced in similar place of articulation as one phoneme. e.g. always perceive [t] as [d] with disregard of vowel length and word position in general. The pseudowords in this condition contain the target voiced phonemes [b], [d], [g], and other voiceless phonemes that share the place of articulation with,

in different word positions (Initial / Middle) (see Table 4 & 5). Since the aim of this paper is investigating the identification of the target phonemes in word final. A separate condition was designed, and there was no need to look at the subjects' ability in identifying the target phoneme in word-final position in this condition. As in the first condition, the native should identify the differences between the phonemes in the middle and initial position, if they could not that means that the stimuli is invalid.

Table 4. Pseudo words for voice and voiceless stops identification in word-initial

/dɛfmæs/	/gonɜl/	/pɛsfæn/	/tæsfʊl/	/kæslim/	/bɔrsan/
/tɛfmæs/	/konɜl/	/bɛsfæn/	/dæsfʊl/	/gæslim/	/pɔrsan/

Table 5. Pseudo words for voice and voiceless stops identification in word middle

/næsdæ/	/heskim/	/nɔfpæs/	/sæmtæn/	/nægmæ/	/sibal/
/næstæ/	/hesgim/	/nɔfbæs/	/sæmdæn/	/nækmæ/	/sipal/

3. Results and Discussion

The obtained data results is discussed in relation to the theoretical and empirical literature. The researcher was interested in assessing the possibility of forming L2 phonetic category beside L1 and the role of language experience. In order to do so, the identification of final stop based on the preceding vowel length was investigated in the perception of Adult L2 learners. For this purpose, two questions were formulated to serve as the research questions. The answers of these two questions were fairly helpful in illustrating the systematic processes by which L2 learners were able to identify final stop voicing.

The results represent subjects' performance in one task with four blocks comprising three controlled conditions followed by one test condition. The controlled conditions aimed to assess the subjects' capability in identifying vowel length differences, and phonemic identification for targeted stops in word-initial/middle position prior to taking the test condition. The test condition was about the identification of voicing for stops in coda position based on the preceding vowel length. In general, the results of all four blocks of the task revealed significant performance differences among the subjects according to on their proficiency levels. It seemed that as their level of proficiency increased, their performance in all four blocks of the task increased (see Figure 3).

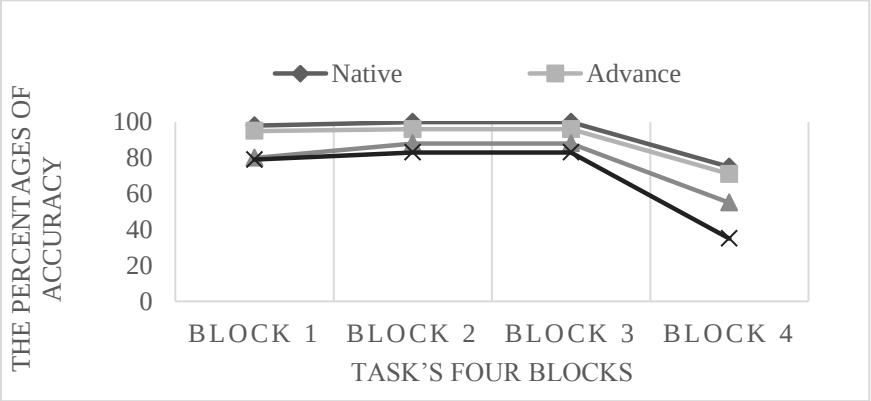

Figure 3. Subjects' overall performance in four blocks of one Task

The one-way ANOVA test showed significant performance among the subjects with different proficiency levels in each block. The data was processed through one-way ANOVA between-subjects, with proficiency level as independent variable (advance, intermediate, and beginners vs native speakers) and each one of the four earlier mentioned blocks in the task as a dependent measure. The initial alpha level for all statistical tests of significance was set at alpha = .05

The first block was about the identification of vowel length, there was a notable effect of subjects' proficiency level in general, F (3, 8) = 12.4, p < .002, but between subjects the post hic test (Bonferroni) revealed that the accuracy scores were not always significant between subjects. There was no significance between (native, advance, & intermediate) and between (intermediate & beginners), but there were significant differences between (native & beginners) and between (advance & beginners), see Table 6 for p value significance between subjects. The accuracy scores out of 24 were as follow: for beginner subjects (M = 19.6), intermediate subjects (M = 21.6), advance subjects (M = 23), and native subjects (M = 23.6).

It is clear from the results that the performance of the subjects varied and improved as their proficiency level of L2 increased. Although Arabic has long and short vowels (Al-Ani, 1970), which means that the subjects could perform well in vowel length distinction, the beginner subjects along with intermediate seemed confused in some words of the stimuli and showed inconsistent errors. This was shown in % 79 success beginners and % 80 success for intermediate. This could be attributed to the nature of the stimuli as it has pseudowords or to the length of the stimuli, which could also be another factor. On the other hand, the native and advanced subjects had a better performance rate % 98 success for native and % 95 success for advance. Overall the results revealed that the subjects generally were capable to differentiate between short and long vowels that could be used as a cue in identifying final stop voicing.

Table 6. p vale significances between proficiency levels in the identification of vowel length

	Native	Advance	Intermediate	Beginner
Native	-	p = 1.0	p = .133	p<.003
Advance	-	-	p = .60	p <.009
Intermediate	-	-	-	p = .133
Beginners	-	-	-	-

The second block was about the identification of English stops in initial position of the word. There was a considerable effect of subjects' proficiency level in general, F (3, 8) = 15.1, p < .001. Bonferroni test revealed that the accuracy scores were not always significant between subjects. There was no significance between (native & advance), (advance & intermediate), and between (intermediate & beginner), but there were significant differences between (native, intermediate & beginner) and between (advance & beginners), see Table 7 for p value significance between subjects. The accuracy scores out of 12 was as follow: for beginner subjects (M = 10), intermediate subjects (M = 10.6), advance subjects (M = 11.6), and native subjects (M = 12).

The subjects' performance in this block also varied between each other. Their performance differed slightly when they were close in proficiency level, but it varies largely as the gap between their levels of proficiency becomes big. Specifically, between subjects, their success rate was as follows: native %100 success, advance %96 success, intermediate % 88 success, and beginners %83 success. When the types of errors were examined, it was found that mostly in words that have the phoneme /b/ and /p/. These two phonemes were expected to be problematic for the Arabic subject as the phoneme /p/ does not exist in Arabic and would cause confusion for the Arabic English learners (Flege, 1980). Similar

results were found in the following block regarding the identification of English stops in the word-middle position.

Table 7. p vale significances between proficiency levels in the identification of stops in word-initial

	Native	Advance	Intermediate	Beginner
Native	-	*P* = 1.0	*P* < .02	*P* <.002
Advance	-	-	*P* = .10	*P* <.006
Intermediate	-	-	-	*P* = .50
Beginners	-	-	-	-

The third block was about the identification of English stops in word middle. Also, there was a significant impact of subjects' proficiency level in general, F (3, 8) = 15.1, p < .001. The planned comparisons (Bonferroni-corrected) revealed similar results that were in the second condition. The accuracy scores were not always significant between subjects. There was no significance between (native & advance), (advance & intermediate), and between (intermediate & beginner), but there was significant differences between (native, intermediate & beginner) and between (advance & beginners), see Table 8 for *p* value significance between subjects. The accuracy scores out of 12 were as follows: for beginner subjects (M = 10), intermediate subjects (M = 10.6), advance subjects (M = 11.6), and native subjects (M = 12). The success rate was as follows: native %100 success, advance %96 success, intermediate % 88 success, and beginners %83 success.

Table 8. p vale significances between proficiency levels

	Native	Advance	Advance	Beginner
Native	-	*P* = 1.0	*P* < .02	*P* <.002
Advance	-	-	*P* = .10	*P* <.006
Intermediate	-	-	-	*P* = .50
Beginners	-	-	-	-

Based on the results obtained from the first 3 controlled conditions, it is possible to say that generally, the subjects seemed to have the knowledge regarding the vowel length distinction and phonemic distinction of English stops (word-initial & middle position). These results indirectly support earlier predictions regarding creating a new phonetic category besides L1 in the underlying phonological system of the subjects and the effect of previous linguistic experience in L2 performance. This would lead us to conclude that in the test condition they may employ vowel length as a key in identifying targeted final stop voicing, along with a potential variation on the subjects' performance, as similarly happened in controlled condition.

The findings of the test condition directly answer the two previously formulated former questions of this paper. The first question in this paper was asked to examine whether the subjects are able to use vowel length that precedes the targeted stops, as a cue in identifying voicing in English. Table 9 demonstrates a summary that provide descriptive statistics for the performance of the subjects on the test condition regarding the identification of final stop voicing.

Table 9. Descriptive statistics for the performance of the subjects on the test condition

Proficiency level	N	Mean SD	%	95% CI		Min	Max
			success	Lower	Upper		
Native	3	9.00	75%	6.52	11.50	8.00	10.00
Advance	3	8.66	72%	4.90	12.50	7.00	10.00
Intermediate	3	6.66	55%	2.90	10.50	5.00	8.00
Beginners	3	4.33	35%	.55	8.12	3.00	6.00

Note: % success was reached through dividing Mean by the overall potential score of 12.

The test condition results were further analysed through a one-way ANOVA along with descriptive statistics to find out whether the level of proficiency is a factor in utilizing the vowel length as a key in identifying voicing on stop. Looking at the distribution of the results within each proficiency level, the assumption of normality has been found cross the distribution of scores (see Figure 4).

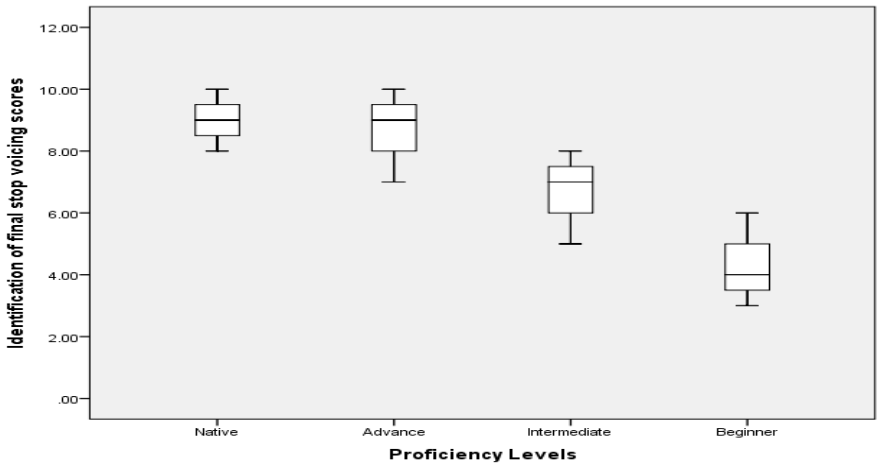

Figure 4. Box plots for subjects' performance in condition test according to proficiency level

Before performing the one-way ANOVA analysis, Levine's test was conducted to examine possible assumption violation for homogeneous variances. The gained outcomes demonstrate that the assumption was detected, F (3, 8) = .33, p = .802. Therefore, the results out of ANOVA is valid and can be interpreted.

According to the employed ANOVA test, the independent variable, which is proficiency level had a notable impact on the subjects' overall performance, in regards to identifying voicing of final stop, F (3, 8) = 6.9, p < .013. This answers the second question concerning the impact of proficiency level development on the establishment of L2 phonetic category. Based on the descriptive statistics, there were an increase in Mean scores from the beginner group and upward (Table 9). The obtained ANOVA data can be translated that the overall ability of the subjects in identifying voicing of final stop in L2 English, their performance improved considerably with their level of competence.

The results showed that the two higher proficiency groups (native & advance) were not notably distinct, when compared to each other, (p = 1.0), because both had close scores on success rates and Mean on the test condition (see Table 9). Likewise, the two groups provide an interesting better performance in identify voicing in stops than the rest of the two lower proficiency groups, as shown by the planned comparisons (Bonferroni-corrected) post hoc analysis (see Table 9), and the box plot of the groups' scores (see Figure 4).

There were no noteworthy variances between (native, advance & intermediate) groups, (advance & intermediate) groups, and between (intermediate & beginner) groups, but there were a high significant differences between (native, & beginner) groups, and (advance & beginners) groups, see Table 10 for p value significance between subjects. The accuracy scores out of 12 was as follows: for beginner subjects (M = 4.33), intermediate subjects (M = 6.66), advance subjects (M = 8.66), and native subjects (M = 9). The success rate was as follows: native %75 success, advance %72 success, intermediate % 55 success, and beginners %35 success.

Table 10. *P* vale significances between proficiency levels

	Native	Advance	Intermediate	Beginner
Native	-	$P = 1.0$	$P = .5$	$P < .022$
Advance	-	-	$P = .7$	$P < .034$
Intermediate	-	-	-	$P = .5$
Beginners	-	-	-	-

Overall, the Bonferroni results revealed that with some exceptions, the level of proficiency is significantly associated with the subjects' ability in identifying voicing of final stop in English. These findings come in favour and support the hypothesis of this study, which stated that if a group of subjects with different proficiency levels in L2 was compared in identification task for final stop voicing in L2 English, then the group with higher proficiency level will outperform the other group with a lower proficiency level due to the progress of L2 phonological category that corresponds to the development of L2 proficiency.

These findings provide some support for the Speech Learning Model SLM (Flege, 1995), which postulates that increase in language experience in L2 may lead to a better perception for L2 sounds as a consequence of creating a new phonetic category. The findings may also be considered as evidence in support of the Dual Language System Hypothesis DLSH (Genesee, 1989), which propose the ability to create a different phonological category for L2 beside L1. In the light of these two hypotheses and obtained results, the bilingual adult subjects' underlying phonological system would be divided into two phonetic categories L1 and L2. Their L2 phonetic category development seemed to parallel with the subject's proficiency development, as illustrated on the subject's performance. This indicates the existence of L2 phonetic category and shows the positive consequence of language experience for the growth of L2 phonetic category.

Regardless of the statistic that the advanced subject performance was not like as a native speaker, they were very close in identifying voicing of a final stop based on the previous vowel duration, native subjects 75% success, and advance subjects 72% success 3% difference between the two. The outcomes substantiate the conclusions of many of earlier studies in the literature regarding the possible establishment of L2 phonetic category beside L1 category (Flege, 1980; Flege & Port, 1981; Fokes et al., 1985; Flege & Efting, 1987). It also comes in line with PAM hypothesis (Best, 1994; Best 1995), and with NLM hypothesis (Kuhl, 2000), clearly in the performance of beginners more than intermediate subjects.

With a 3 % different slight performance between advance and native subjects, it is possible to say that these results contradict those reported by Flege et al., (1995) that bilinguals' performance in L2 perception is not close similar to native speakers' perception. They also contradict the results reported in the literature for other study (Darcy & Krüger, 2012) regarding the influence of the L1's phonetic category upon the L2's phonetic category in the perception of the L2. Neither does it align with the conclusions reported by (Ohde & Abou-Khalil, 2001), concerning the influence of age on bilinguals' L2 perception.

Based on these outcomes, one can postulate that the bilingual advanced subjects had established two phonetic categories in their underlying phonological systems. These two phonetic categories are L1 Arabic and L2 English phonetic categories. During the perception process these two categories could function separately from each other with limited possible influence as found in other studies (Flege, 1980; Flege & Efting, 1987). Although Arabic has no voicing dependent distinction for stops (Flege 1979), it seemed that the advance subjects have succeeded in acquiring L2 specific acoustic feature, vowel length, utilized as a cue for phonetic contrast between voice and nonvoiced stops in English, which helped them to use the same cue for English pseudowords. According to McAllister et

al., (2002), the L2's lexical properties push the phonological system to establish a phonetic L2 category that is separate from the L1 in L2 acquisition. This means that based on the subject proficiency level and lexical knowledge, the English-like pseudowords were processed through one of two existing phonetic categories, which in our case here seemed to be English.

4. Conclusion

To sum it up, the hypothesis of this study was born out, it appeared that the L2 adults have formed a new phonological category for L2 beside their L1, which helped them to identify final stop voicing in English based on the succeeding vowel length as a cue. The establishment of L2 phonetic category seemed to correspond to the improvement of the L2 proficiency level. In Overall, with a 3 % difference, the advance subjects' performance was found to be similar to native subjects. Out of this study, there are some recommendations to be made. Although it was not easy to form English pseudowords, it is suggested that it would be more beneficial to form more English-like words and increase the number of words in the stimuli. Looking at the number of subjects, the outcomes of this study cannot be generalized, and it is recommended to use a higher number of subjects to be able to generalize the findings. Also, duplicating this study with similar subjects, but with different L1 will add to the current discoveries and may lead to have common ground in supporting this paper's hypothesis.

6. Ethics Committee Approval

The author(s) confirms(s) that the study does not need ethics committee approval according to the research integrity rules in their country (Date of confirmation: 12/30/ 2021)

References

Al-Ani, S. (1970). *Arabic Phonology: an acoustical and phonological investigation.* The Hague, Paris: Mouton de Gruyter publisher.

Aljaser, F, Jackson, T. K., Vitevitch, M. S., & Joan A. S. (2018). The influence of phonemeinventory on elicited speech error in Arabic speakers of English. *The Mental Lexicon, 13*, 26-37.

Al Malwi, M. I. (2017). *The effect of gender and age on voice onset time by Abah Arabicspeaker.* (Unpublished master's thesis) California State University, USA.

Alotaibi, A. (2013). *Pronunciation Problem in the production of the voiced labiodentalfricative /v/ by Saudi speaker of English.* (Unpublished master's thesis). Southern Illinois University Carbondale, Carbondale, Illinois, USA.

Alotaibi, A. (2018). *The role of native language dialects on the perception of L2 English vowels.* (Unpublished doctoral dissertation). Indiana University at Bloomington, USA.

Best, C. (1995). A direct realist perspectives on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-Language research* (pp. 171-206). Timonium, MD: York Press.

Boeirsma, P., & Wenink, D. (2017). Praat: doing phonetic by computer [Computer program]. Version 6.0.28, retrieved 23 March 2017 from http://www.praat.org/

Buali, I. (2010). *The Perception and Production of /p/ in Saudi Gulf Arabic English:*

A Variationists Perspectives. (Unpublished master's thesis). Concordia University Montréal, Québec, Canada.

Charles, L. J. (1985). Word-final devoicing in German: Effect of phonetic and sentential contexts. *Journal of Phonetics, 13*, 309–324.

Chen, M. (1970). Vowel length variations as a function of the voicing of the consonants environment. *Phonetica, 22*, 129-159.

Darcy, I., & Krüger, F. (2012). Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetic, 40*, 568-581.

Flege, J. E. & Eefting, W. (1987). Production and perception of English stop by native Spanish speakers. *Journal of Phonetics, 15*, 67-83.

Flege, J. E. (1979). *Phonetic interferences in second language acquisition* (Unpublished doctoral dissertation). Indiana University, Bloomington.

Flege, J. E. (1980). Phonetic approximations in second language acquisition. *Language Learning, 30*, 117-134.

Flege, J. E., Hllengbrand, J., Ingrisano, R., & Smth, B., (1984). Perception of the voiced voiceless contrasts in syllable-final stops. *Journal of the Acoustical Society of America, 76*, 18–26.

Flege, J. E., McCutcheon, M. & Smith, S. (1987). The development of skills in producing word-final English stop. *Journal of the Acoustical Society of America, 82*(2), 433-447.

Flege, J. E. & Port, R. (1981). Cross-language phonetic interferences: Arabic to English. *Language and Speech, 24*, 125-146.

Flege, J. E. (1995). Second language speech learning: Theory, finding, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross language research* (pp. 233-276). Timonium, MD: York Press.

Flege, J. E., Munro, J., & MacKay, I. (1995b). Effect of age of second-language learning on the production of English consonants. *Speech Communication, 16*, 1-26.

Fokes, J., Bond, Z. & Steinberg, M. (1985) Acquisition of the English voicing contrast by Arab children, *Language and Speech, 28*, 81-91.

Genesee, F. (1989). Early bilinguals development: One language or two?. *Journal of Child Language, 16*, 161-179.

Hazan, V. L. & Boulakia, G. (1993). Perception and production of a voicing contrasts by French-English bilingual. *Language and Speech, 36*(1), 17–38.

Kuhl, P. (2000). A new view of language acquisition. *Proceeding of the National Academy of Science 97*, 11850-11857.

Krause, S. (1982). "Developmental use of vowel duration as a cue to postvocalic stops consonant voicing. *Journal of Speech, Language, and Hearing Research, 25*, 388-393.

Ladefoged, P. (2006). *A course in phonetic (5th ed.).* Boston: Thomson Wadsworth.

Lehman, M., & Sharf, D. (1989). "Perception/production relationship in the development of the vowel duration cue to final consonant voicing. *Journal of Speech, Language, and Hearing Research, 32*, 803–815.

Mack, M. (1982). Voicing-dependent vowel durations in English and French: monolingual and bilingual production. *Journal of the Acoustical Society of America, 71,* 173-78.

Major, R. (1987). The natural phonology of second language acquisition. In James, A.and Leather, J. (Eds.), *Sound patterns in second language acquisition*, (pp.207-224). Dordrecht Holland: Foris Publications.

McAllister, R., Flege, J, & Piske, T (2002). The influence of the L1 on the acquisition of Swedish vowel quantity by native speaker of Spanish, English, and Estonian. *Journal of Phonetic, 30,* 229-258.

Mayo, C., & Turk, A. (2004). Adult–child differences in acoustics cue weighting are influenced by segmental context: Children are not always perceptually biased toward transitions. *Journal of the Acoustical Society of America, 115,* 3184–3194.

Mora, J. C. (2007). Learning context effects on the acquisition of a second langua phonology. In C.Pérez-Vidal, M. Juan-Garau, & A. Bel (Eds.), *A portrait of the young in the new multilingual Spain* (pp. 241-263). Clevedon: Multilingual Matter.

Ohde, R. N., & Abou-Khalil, R. (2001). Age differences for stops consonant and vowel perception in adults. *Journal of the Acoustical Society of America, 110,* 2156– 2166.

Port, R., Al-Ani, S., & Maeda, S. (1980). Temporal compensations and universal phonetics. *Phonetica, 37,* 235-252.

Smith, B., Hayes-Harb, R., Bruss, M., & Harker, A. (2009). Production and perception of voicing and devoicing in similar German and English word pair by native speakers of German. *Journal of Phonetics, 37,* 257–275.

Trofimovich, P. (2008). What do second language listener know about spoken words? Effects of experience and attention in spoken word processing. *Journal of Psycholinguistic Research, 37,* 309–329.

Watson, J. C. (2002). *The Phonology and Morphology of Arabic*. Oxford University Press Inc. New York.

Zampini, M. L. & Green, K. P. (2001). The voicing contrasts in English and Spanish: the relationship between perception and production. In Nicol, J. L. (Ed.), *One mind, two languages* (pp. 23–48). Oxford: Blackwell.

AUTHOR BIODATA

Dr. Abdullah Nijr Al-Otaibi is an Assistant professor of Applied Linguistics, Department of English, College of Science and Humanities in Rumah, Majmaah University, Al-Majmaah, 11952, Saudi Arabia. His research interests are second language Learning, E-learning, L2 phonetics and phonology, and developmental speech perception.