First-Language Phonetic Drift of Yemeni Arabic Stops

Sumaya Faisal Alshamiri [a]
alshamiri99@yahoo.com
Center for Research in Language and Linguistics,
Faculty of Social Sciences & Humanities,
Universiti Kebangsaan Malaysia, Malaysia

Shahidi A.H. [b]
zedic@ukm.edu.my
Center for Research in Malay Language, Literature and Culture,
Faculty of Social Sciences & Humanities,
Universiti Kebangsaan Malaysia, Malaysia

Sharifah Raihan Syed Jaafar
s_raihan@ukm.edu.my
Center for Research in Language and Linguistics,
Faculty of Social Sciences & Humanities,
Universiti Kebangsaan Malaysia, Malaysia

ABSTRACT

This study aims at examining the occurrence and direction of any phonetic interlanguage interference among Yemeni Arabic-English bilinguals by indicating the acoustic similarities and differences of L1 Yemeni Arabic stops produced by bilinguals and comparing it to monolinguals. It investigates Voice Onset Time of stops in initial and medial position as well as Preceding Vowel Duration of stops in medial position. A total of 60 native Yemeni Arabic subjects were involved in this study. Thirty subjects were late bilinguals with English as their L2 whereas the other 30 were monolinguals. Data was collected through two production tests: one for the bilingual group and the other for the monolingual group. All the subjects were asked to produce a list of Arabic words with the target stops /b, t, d, k/ in word-initial and word-medial position. Each subject in both samples was recorded individually. The bilingual subjects showed significantly longer mean Voice Onset Time values than monolinguals for /t, k/ whereas they showed mean Voice Onset Time values for the voiced stops /b, d/ that were very close in values to the monolingual group. However, bilinguals showed shorter mean values than monolinguals for the Preceding Vowel Duration, but the difference was not significant. The findings reveal that bilinguals showed signs of English-induced L1 phonetic drift in their pronunciation of Arabic /t, k/ whereas not in their realisation of Arabic /b, d/. This study provides conclusive evidence that L2 characteristics can systematically affect matured L1 phonological systems even among late bilinguals.

Keywords: acoustic phonetics; Arabic-English bilingual; L1 drift; stops; Voice Onset Time (VOT)

---

[a] *Main author*
[b] *Corresponding author*

INTRODUCTION

A growing body of research has proved the influence of L2 on L1 at the phonetic level (Bergmann et al., 2016; Chang, 2012; Dmitrieva et al., 2020; Mayr et al., 2012, 2020; Mora et al., 2015; Osborne & Simonet, 2021; Schwartz, 2020). In the Speech Learning Model (SLM), Flege (1995) postulates that "phonetic categories in childhood for L1 sounds evolve over the lifespan to reflect the properties of all L1 and L2 phones" (p. 239). Studies investigating the phonetic production of bilinguals (Kartushina et al., 2016; Mora et al., 2015) support this claim. Furthermore, L2 learning can lead to three possible effects on native categories: "(1) no change, (2) drift toward the L2 category and (3) deflection away from the L2 category, to maximize opposition with it" (Kartushina et al., 2016, p. 22). Some studies refer to the phonetic effect of L2 on L1 as L1 drift (Chang, 2012; Dmitrieva et al., 2020; Mayr et al., 2020; Osborne & Simonet, 2021).

The Bilingual Category Hypothesis is a SLM hypothesis tested in the present study. It predicts a bidirectional influence between the L1 and L2 in which L1 and L2 sounds of a bilingual are significantly different from the sounds of monolingual speakers of both languages. When L2 sounds are perceived as being phonetically different from the closest native category, new categories are expected to be established for them. Dissimilation of a phonetic category takes place when bilinguals develop a distinct category for the newly encountered L2 sounds. Assuming that the L1 and L2 phonetic categories coexist within the same perceptual acoustical space, perceptually similar L1 and L2 categories may drift apart since they are sharing a more crowded space (Flege et al., 2003). The acquisition of a similar L2 sound may cause the two categories to drift to avoid crowding the phonetic space.

To investigate the presence of any L2 effect on L1, a comparison of bilingual speech and monolingual speech is essential. Besides, when bilinguals produce monolingual-like speech for both their languages, it is an indication that they have developed separate language-specific categories. On the other hand, if their L1 speech differs from that of monolinguals, that is taken as evidence for an L2 effect on L1. In the present study, the possibility of a cross phonetic interaction between L1 Arabic – L2 English is examined. It is hypothesised that the bilingual participants produce sounds that are different from those produced by Arabic monolinguals.

Several studies have demonstrated that late acquired L2 phonology influences L1. In addition, studies have reported that L1 phonetic systems may drift towards L2 phonetic categories in late bilinguals who have been exposed to an L2 environment (Bergmann et al., 2016; Chang, 2012). For instance, Chang (2012) investigated the Korean of native English Learners enrolled in a Korean intensive language course. In a word reading task, the VOT values of English voiceless stops were drifted toward Korean. Besides, L1 drift towards L2 may as well occur in early bilinguals. For example, Mackay et al. (2001) observed the voiced stops of early and late Italian-English bilinguals. The Italian production of both bilingual groups differed from that of Italian monolinguals. Since /b, d, g/ are typically produced with short-lag VOT values in English whereas they are produced with lead VOT values (prevoicing) in Italian, bilinguals misidentified the English /b, d, g/ stops as /p, t, k/. Their English voiced stops were produced with lead VOT values more often than that of native English speakers whereas their Italian voiced stops were produced with lead VOT values less often than that of Italian monolinguals. Precisely, the Italian /b/ was produced with lead VOT values less often than Italian monolinguals. Furthermore, drift of L1 towards L2 categories has also been reported among early bilinguals whose dominant language was their L2 (Mora et al., 2015).

Changes in phone production of a bilingual is not always assimilatory in nature. The influence of L2 on L1 sounds may as well appear in the form of L1 sounds drifting away from the L2. During production, speakers create contrast between similar sounds in a common L1-L2 phonological space by exaggerating their dissimilarities. In other words, an L1 category may drift away from that of monolingual speakers of L1 to maintain contrast with the similar L2 phone category. Based on the SLM, L1 drifts away from L2 so that both categories are kept maximally distinct (Flege, 1995). For instance, Harada (2003) investigated the stop production of early Japanese-English bilinguals. Compared to Japanese monolinguals, these bilinguals' Japanese voiceless stops were produced with longer VOT values whereas their English voiceless stops had native-like VOT values. These findings imply that to maintain phonetic contrast in a common phonological space, it is L1 sounds and not L2 sounds that can be deviated from L1 phonetic categories (Harada, 2003). Overall, the findings of these studies indicate that L2 may influence L1 by either drifting the L1 toward or away from it. In other words, cross-linguistic interference may not necessarily result in assimilation of L1 and L2 categories but instead may lead to dissimilation of these categories. Such an effect of L2 on L1 has been observed mostly among bilinguals who are L2 proficient or among those who are immersed in an L2 setting (Kartushina et al., 2016).

Conflicting findings exist as whether bilinguals can produce VOT values equivalent to monolingual speakers in each of their languages. Some studies reported that VOTs of bilingual speakers failed to match those of monolingual speakers of either one or both languages (Harada, 2003; Schwartz, 2020). However, Antoniou et al. (2020) reported that VOTs of bilinguals matched those of monolinguals in both languages in which, for the two languages of the bilinguals, the voiced and voiceless categories were not equally spaced along the VOT dimension. In other words, previous research suggests that bilingual speakers distinguish between VOT values among their L1 and L2 to varying degrees. While some bilinguals have a single phonetic category with intermediate VOT values, others hold distinctly different VOTs which may also differ from monolingual speakers of those two languages. This difference is an indication of two distinct phonetic categories of a bilingual's L1 and L2.

The current study examined VOT production in the L1 Arabic of late Yemeni Arabic-English bilinguals. It studies the L1 of a language pair that has not been investigated before. It is more common that Arabic has a voicing opposition between voicing lead and short lag, while in English the distinction is between short lag and long lag. Given that the phonetic properties of related L2 categories can influence L1 phonological categories, this study investigated the occurrence and direction of any phonetic interlanguage interference among Yemeni Arabic-English bilinguals. This was fulfilled by examining the acoustic similarities and differences of L1 Yemeni Arabic stops (Taizzi dialect) produced by bilinguals and by comparing it to monolingual production. These objectives were fulfilled through answering the following research questions:

1. What are the acoustic similarities and differences of Taizzi Yemeni Arabic stops produced by bilingual and monolingual Yemeni subjects?
2. Is there any phonetic interlanguage interference among Taizzi Yemeni Arabic-English bilinguals?

ARABIC AND ENGLISH VOT VALUES

Within the literature on second language acquisition, VOT as introduced by Lisker and Abramson (1964) has been widely investigated due to its precise nature and its language specific variation. Any description of a language's phonetic structures should include an observation of the VOT since it varies considerably across languages (Ladefoged, 2003). Thus, it is one of the acoustic measurements under investigation in this study. Lisker and Abramson (1964) define VOT as the "interval between the release of the stop and the onset of glottal vibration" (p. 389). They point out that if the vibration of the vocal cords starts before the closure release, the VOT value is negative "voice lead" whereas if the vibration starts after the closure release, the VOT value is positive "voice lag".

The literature reveals that there is indeed a lot of language-specific diversity when it comes to VOT. Several languages follow a voicing lead and lag VOT pattern such as Italian (Mackay et al., 2001), Malay (Shahidi & Aman, 2011), Polish (Schwartz, 2020), Portuguese (Osborne & Simonet, 2021), and Spanish (Tobin et al., 2017), whereas other languages follow a short lag and long lag pattern such as English (Docherty, 1992) and Japanese (Harada, 2003). In addition, there are different varieties of Arabic depending on the region/country where it is spoken. Some of these varieties are mutually unintelligible to other Arabs who are unfamiliar with the dialect. Due to obvious dialectal variations of different Arabic regions features are to vary accordingly. Based on their findings, Fox and Jacewicz (2009) revealed that regional dialects are a rich source of phonetic variation. Several studies examined the Arabic dialects spoken in Yemen, Saudi Arabia, Lebanon, Qatar, Jordan, Iraq (Mosul), Egypt, Palestine (Al-Nuzaili, 1993; Flege & Port, 1981; Khattab, 2002; Kulikov, 2018; Mitleb, 2009; Rahim & Kasim, 2009; Rifaat, 2003; Tamim & Hamann, 2021), respectively. According to the literature, acoustic values vary from one Arabic dialect to another. Arabic does not follow one single pattern. VOT of Arabic stops varies according to the different Arabic dialects. Some studies revealed that VOT values are always positive, thus following the pattern of short lag and long lag for voiced and voiceless stops, respectively (Aldahri, 2012; Aldahri & Alotaibi, 2010; Flege & Port, 1981; Mitleb, 2009). On the other hand, other studies on specific Arabic dialects revealed that the VOT values are negative (voicing lead) for voiced sounds and positive for voiceless sounds (Adam, 2012; Al-Nuzaili, 1993; Alsiraih, 2020; Khattab, 2002; Kulikov, 2018; Rahim & Kasim, 2009; Rifaat, 2003; Tamim & Hamann, 2021; Yeni-Komshian, Grace H. Caramazza & Preston, 1977).

Al-Nuzaili (1993) investigated the emphasis and voicing distinction of stop production of Yemeni Arabic. He observed the VOT distinction of stop consonants. He also examined whether the vocalic context affects the VOT value of the preceding stop consonant. The Arabic stops /b, t, tˤ, d, dˤ k, g, q/ in word-initial position were examined. All were put in monosyllabic words followed by these six vowels; /i, a, u, i:, a:, u:/. Recordings were collected from only one subject. Results showed that voiced stops of Yemeni Arabic have their VOT values in the voice lead region whereas the voiceless ones occupied the voicing lag. Unlike other studies on Arabic which follow the voicing lead vs. lag pattern, VOT ranges for /t, k/ exceed the short lag period determined by Lisker and Abramson (1964). On the other hand, some voiced sounds were associated with a long voicing lead reaching a maximum value of -120 msec for /b/ and -130 msec for /d, g/. Table 1 below summarises VOT patterns of several Arabic dialects.

eISSN: 2550-2131
ISSN: 1675-8021

TABLE 1. Studies on VOT Patterns of Arabic Dialects

Arabic VOT Studies	**Arabic Dialect Investigated**	**VOT Pattern (voiced vs. voiceless)**
Yeni-Komshian, Caramazza & Preston (1977)	Lebanese	voice lead vs. short lag
Al-Nuzaili (1993)	Yemeni	voice lead vs. lag
Khattab (2002)	Lebanese	voice lead vs. short lag
Rifaat (2003)	Egyptian	voice lead vs. short lag
Mitleb (2009)	Jordanian	short lag vs. long lag
Rahim & Kasim (2009)	Iraqi	voice lead vs. short lag
Aldahri & Alotaibi (2010)	Saudi Arabian, Jordanian & Modern Standard Arabic	short lag vs. long lag
Aldahri (2012)	Modern Standard & Classical Arabic	short lag vs. long lag
Adam (2012)	Palestinian	voice lead vs. short lag
Kulikov (2018)	Qatari	voice lead vs. short lag
Alsiraih (2020)	Iraqi	voice lead vs. short lag
Tamim & Hamann (2021)	Palestinian	voice lead vs. short lag

Furthermore, several studies investigated VOT values of native English dialects showing the pattern short lag and long lag for voiced and voiceless stops, respectively (Antoniou et al., 2010; Docherty, 1992; Lisker & Abramson, 1964). Besides, numerous research has been conducted on the VOT of particular dialects of English (Antoniou et al., 2010; Chang, 2012; Flege & Port, 1981; Naji, 2019; Shahidi & Aman, 2011; Tobin et al., 2017). One research which is the most relevant to this study is that of Naji (2019). He examined English stops produced by native Yemeni Arabic speakers of English and then compared them to VOT values of native English subjects presented in the literature. Moreover, he stated that his aim is to identify the effect of Yemeni speakers' L1 Arabic on their production of L2 English stops. However, there was no investigation made on their L1 Arabic counterparts.

Besides, though Al-Nuzaili (1993) investigated L1 Yemeni Arabic, an investigation on L1 Yemeni Arabic produced by Arabic-English bilinguals and comparing it to Yemeni Arabic produced by monolinguals is essential. The present study goes beyond previous work in that it examines the acoustic features of L1 Yemeni Arabic produced by bilinguals and compares it to monolinguals to observe any L2 influence on L1. Unlike Al-Nuzaili (1993) where speech samples were collected from only one subject, the present study has an adequate number of subjects involved. This study provides extensive findings on the phonetic properties of a variety of Arabic spoken by Yemeni speakers.

METHOD

PARTICIPANTS OF THE STUDY

The dataset included 60 participants altogether, all of whom are of Yemeni origin. There were 30 monolingual and 30 bilingual subjects. Of the monolingual participants, 16 were female and 14 were male, aged between 25 and 54 years old, whereas of the bilingual subjects, 17 were female and 13 were male, aged between 27 to 47 years old. All the participants spoke the Taizzi dialect, which is a variety of Yemeni Arabic (YA) spoken by a large population. Overall, there are four major regional dialects in Yemen: Sana'ani, Taizzi, Tihami and Hadhrami dialects. The Taizzi dialect covers the region of Taiz, Ibb, and Aden (Salem & Pillai, 2020). All participants had their primary and secondary school education in Yemen. They have all acquired Modern Standard Arabic at school which is the formal written form of Arabic.

All participants are residents of Klang Valley, Malaysia. They live within a Yemeni community which resembles a strong Yemeni environment such as Yemeni shops, restaurants and other services, therefore, exposure to other languages (if any) is very scarce. Two language background questionnaires were designed. One was designed for the monolingual sample and one for the bilingual sample. The aim of the questionnaire was to check that each subject is eligible to be part of the sample. Once the target number of subjects fulfilled the selection criteria, the questionnaire was no longer distributed. A different selection criterion was followed for each sample.

For the monolingual sample, any subject who reported acquiring a language other than their mother tongue and/or has lived in Malaysia for more than 6 years was not part of the sample. Furthermore, they all have had little if any exposure to any foreign language. They manage to live without needing to communicate in English nor Malay. They have received their basic education in Yemen. Eight reported not finishing their secondary education while six have finished their secondary education. Besides, 14 reported having an undergraduate degree and 2 having postgraduate degrees.

For the bilingual sample, all subjects reported speaking Yemeni Arabic (Taizzi dialect) as their L1 and English as L2. They reported being fluent in both languages and acquire no other language. Moreover, their L1 was acquired since birth whereas their L2 was mastered after graduating from secondary school. All the bilingual subjects are postgraduate students specialising in various fields, with English being their medium of instruction. All subjects are students at three public universities in Klang Valley, Malaysia. These universities are University of Malaya (UM), Universiti Kebangsaan Malaysia (UKM), and Universiti Putra Malaysia (UPM). Subjects have reported being residents in Malaysia from 1 to 6 years with a mean average of 4 years.

LINGUISTIC MATERIAL

The linguistic material consisted of an Arabic word list. The linguistic variables investigated are the Arabic stops /b, t, d, k/. For each target stop sound, there were two words on the list, one with the target sound in the initial position and another in the medial position. Moreover, to maintain the speech rate, additional words were added to the beginning and ending of the list.

However, these words were later excluded from the analysis. For the target stop sound in the initial position, the syllable structure is CVC where the syllable nucleus consists of the following vowel /iː/. Additionally, for words with the target stop sound in the medial position, the syllable structure is CVCVCVC. The target consonant sound is the third consonant in the word which is placed in the second syllable and preceded by the vowel /ɪ/.

DATA COLLECTION PROCESS

The data collection process involved two oral production tests in which target sounds were recorded then acoustic cues were measured, therefore, a quantitative research design was followed. The first production test (PT1) was applied on the monolingual sample whereas the second production test (PT2) was applied on the bilingual sample. Each subject was required to read each word on the list five times. The recording session for each subject took place individually. Altogether, 60 recording sessions were held for both production tests.

A Sony IC Recorder model (ICD-UX560F) with a built-in microphone and speaker was used for recording. During the recording session, the researcher held the recorder in her hand around 20 cm away from the subject's mouth. The audio recordings were then labelled and saved. Each speech sample was analysed via Praat software (Boersma & Weenink, 2019). Measurements were made based on the display of the waveform and spectrogram associated with the target sound. Besides, measurements were made manually by positioning two cursors, one indicating the beginning and the other indicating the ending of a particular acoustic measurement. For target stops in the initial position, VOT was measured whereas for those in the medial position, VOT as well as the preceding vowel duration were investigated in this study.

PT1 investigated the acoustic features of Arabic stops produced by monolingual Yemeni subjects, therefore, samples of Arabic containing the target sounds /b, t, d, k/ were collected. For PT1, there were 1200 consonant tokens for analysis (4 sounds x 2 word-positions 'initial & medial' x 30 subjects x 5 repetitions), a total of 1200 tokens under investigation. Furthermore, for (PT2), samples of Arabic containing the target sounds /b, t, d, k/ were collected from the bilingual subjects. Since both samples are equal and the same Arabic material was used in both tests, PT2 had an equal number of tokens investigated to those of PT1. Trials in which the subject mispronounced the word or hesitated before reading the word were excluded from the analysis (19 tokens from PT1 and 10 tokens from PT2 were excluded). Leaving a total of 1181 monolingual Arabic tokens and 1190 bilingual Arabic tokens. The full dataset included 2371 tokens.

RESULTS

The goal of this analysis is to examine the acoustic properties of native Yemeni Arabic produced by monolinguals and Arabic-English bilinguals. In addition, this analysis determines whether bilinguals' productions of initial and medial Yemeni Arabic stops were affected by acquisition of L2 English. Significant differences in the acoustic realisation in terms of VOT and Preceding Vowel Duration (PVD) of initial and medial Arabic stops between the bilingual and monolingual groups would reveal this impact. In this section, the bilingual group is referred to as (BI) while the monolingual group is referred to as (MO).

VOT IN INITIAL POSITION

An independent sample T-test was conducted on all the target sounds in initial position. For mean VOT values of YA initial /b/, MO (N = 30) showed a mean value and standard deviation (M = -69.51, SD = 23.13) whereas BI (N = 28) showed (M = -68.84, SD = 18.38). The assumption of homogeneity of variance was tested and satisfied via Levene's *F* test, therefore, equal variances were assumed, F (56) = 2.05, p = .158. Since the significant value was larger than alpha, there were no statistically significant differences between initial /b/ produced by MO and BI, t (56) = -.120, p = .905, 95%, CI [-11.70, 10.38].

Additionally, for mean VOT values of /d/, MO (N = 29) showed a mean value and standard deviation (M = -82.49, SD = 21.84) whereas BI (N = 30) showed (M = -83.72, SD = 21.95). The assumption of homogeneity of variance was tested via Levene's *F* test and equal variances were reported, F (57) = .003, p =.955. The independent sample T test has revealed that there were no statistically significant differences, t (57) = .215, p = .830, 95%, CI [-10.19, 12.64]. Thus, there were no statistically significant differences between /d/ produced by MO and BI.

eISSN: 2550-2131
ISSN: 1675-8021

For mean VOT values of /t/, MO (N =30) showed a mean value and standard deviation (M = 48.11, SD = 18.85). By comparison, BI (N = 30) showed a numerically higher mean VOT value (M = 63.37, SD = 12.12). The assumption of homogeneity of variances was violated, therefore, unequal variances were reported, F (49.47) = 5.09, p = .028. With the significant value being less than alpha, the independent sample T test revealed statistically significant differences, t (49.47) = -3.73, p < .000, 95%, CI [-23.48, -7.03] between BI and MO in the production of /t/.

Moreover, mean VOT values of /k/, MO (N =30) showed a mean value and standard deviation (M = 55.92, SD = 12.03). By comparison, the bilingual group (N = 30) showed a numerically higher mean VOT value (M = 68.59, SD = 10.87). The assumption of homogeneity of variance was tested via Levene's *F* test, F (58) = .206, p = .652 and equal variances were assumed. The independent sample T test has revealed statistically significant differences, t (58) = -4.28, p < .000, 95%, CI [-18.60, -6.75]. Thus, BI was associated with a statistically significant larger VOT mean for /t/ and /k/ than MO. Figure 1 and Table 2 below illustrate the mean VOT values in milliseconds of YA stops in word-initial position produced by monolingual and bilingual subjects.

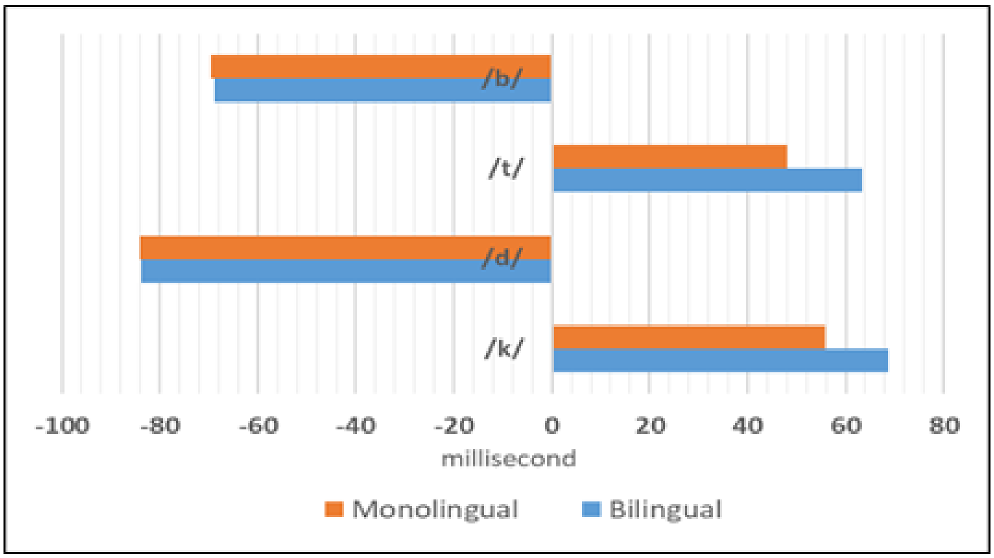

FIGURE 1. Mean VOT values in milliseconds of YA stops in word-initial position produced by monolingual and bilingual subjects

TABLE**Error! No text of specified style in document.** 2. Descriptive and Inferential Statistics for VOT of Arabic stops in word-initial position produced by monolingual and bilingual subjects

Sound	Monolingual		Bilingual		df	T	Sig
	Mean	Std. Deviation	Mean	Std. Deviation			
/b/	-69.50	23.13	-68.84	18.38	56	-0.120	p = .905
/t/	48.11	18.85	63.37	12.12	49.47	-3.728	p = .000
/d/	-84.11	23.21	-83.72	21.95	57	0.215	p = .830
/k/	55.92	12.03	68.59	10.87	58	-4.282	p = .000

VOT IN MEDIAL POSITION

Both MO and BI showed lag VOT values for voiceless stops and lead VOT values for voiced stops. The typical VOT pattern of /b/ and /d/ by both MO and BI showed that voicing in the closure phase overlapped with the prevoicing (voicing lead) of the target stop leaving no voicing gap (break) throughout the stop closure phase, therefore, the stop closure and voicing could not be differentiated. Thus, VOT was not measured for medial /b/ and /d/ tokens. In addition, (33%) of the tokens of medial /b/ by MO showed an inaudibly released burst (i.e., without a stop burst) while (20%) for BI. On the contrary, all medial /d/ produced by MO and BI were audibly released.

Figure 2 and Table 3 below illustrate the mean VOT values in milliseconds of YA stops in word-medial position produced by monolingual and bilingual subjects. For mean VOT values of /t/, MO (N = 30) showed a mean value and standard deviation (M = 26.39, SD = 7.35) whereas BI (N = 29) showed (M = 29.16, SD = 7.29). The assumption of homogeneity of variance was tested and satisfied via Levene's *F* test, F (57) = .469, p = .496. The independent sample T test has revealed that there were no statistically significant differences between MO and BI, t (57) = -1.46, p = .151, 95% CI [-6.60, 1.04]. However, for the mean VOT values of /k/, MO (N =30) showed a mean value and standard deviation (M = 32.59, SD = 6.51) while BI (N = 30) showed (M = 37, SD = 8.47). The assumption of homogeneity of variance was tested via Levene's *F* test, F (58) = 3.42, p = .070 and unequal variances were reported. The independent sample T test has revealed statistically significant differences, t (58) = -2.26, p = .028, 95% CI [-8.31, -.50]. Thus, BI was associated with a statistically significant higher mean VOT for /k/ than MO.

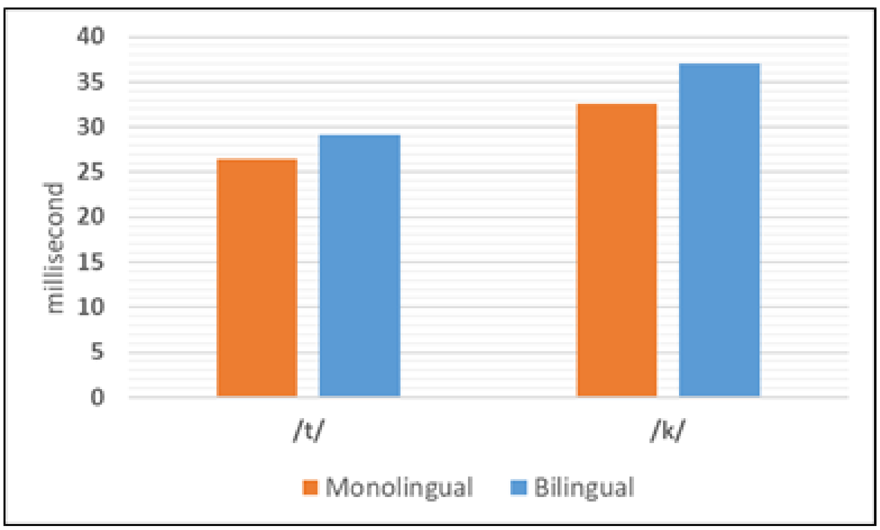

FIGURE 2. Mean VOT values in milliseconds of YA /t/ and /k/ in word-medial position produced by monolingual and bilingual subjects

TABLE 3: Descriptive and Inferential Statistics for VOT of YA word-medial stops /t/ and /k/ produced by monolingual and bilingual subjects

Sound	Monolingual		Bilingual		df	T	Sig
	Mean	Std. Deviation	Mean	Std. Deviation			
/t/	26.39	7.35	29.16	7.29	57	-1.46	p = .151
/k/	32.59	6.51	37.00	8.47	58	-2.26	p = .028

PRECEDING VOWEL DURATION (PVD)

The duration of the preceding vowel was measured by positioning the cursors at the onset of energy in F1 indicating the start of the vowel to the point of a sharp decline in F1 and F2 indicating the end of this vowel. Mean PVD values of /b/ for MO (N = 30) are (M = 45.25, SD = 11.63) while they are (M = 41.39, SD = 11.16) for BI (N = 30). An independent sample T test was conducted. Besides, the assumption of homogeneity of variance was tested and accepted via Levene's *F* test, F (58) = .237, p = .628. Based on the independent sample T test, t (58) = 1.31, p = .195, 95%, CI [-2.03, 9.75], there were no statistically significant differences between PVD values of /b/ by MO and BI.

For /d/, MO (N = 29) showed a mean value and standard deviation (M = 54.90, SD = 11.39) whereas BI (N = 28) showed (M = 49.66, SD = 11.61). The assumption of homogeneity of variance was tested and equal variances were reported, F (55) = .026, p = .872. The independent sample T test has revealed that there were no statistically significant differences between PVD values of /d/ by MO and BI, t (55) = 1.72, p = .091, 95% CI [-.864, 11.34].

In addition, mean PVD values of /t/ for MO (N =30) are (M = 36.63, SD = 11.28) while BI (N = 30) showed (M = 33.14, SD = 9.94). Moreover, the assumption of homogeneity of variance was tested and accepted via Levene's *F* test, F (58) = .050, p = .823. Based on the independent sample T test, t (58) = 1.27, p = .209, 95%, CI [-2.01, 8.99], there were no statistically significant differences between PVD values of /t/ by MO and BI.

Besides, mean PVD values of /k/ for MO (N =30) are (M = 37.96, SD = 11.04) BI (N = 30) showed (M = 35.54, SD = 9.71). The assumption of homogeneity of variance was tested via Levene's *F* test and equal variances were reported, F (58) = .047, p = .829. The independent sample T test has revealed that there were no statistically significant differences between PVD values of /k/ by MO and BI, t (58) = .90, p = .371, 95%, CI [-2.96, 7.79]. Figure 3 and Table 4 below illustrate the mean PVD values in milliseconds of YA stops in word-medial position produced by monolingual and bilingual subjects.

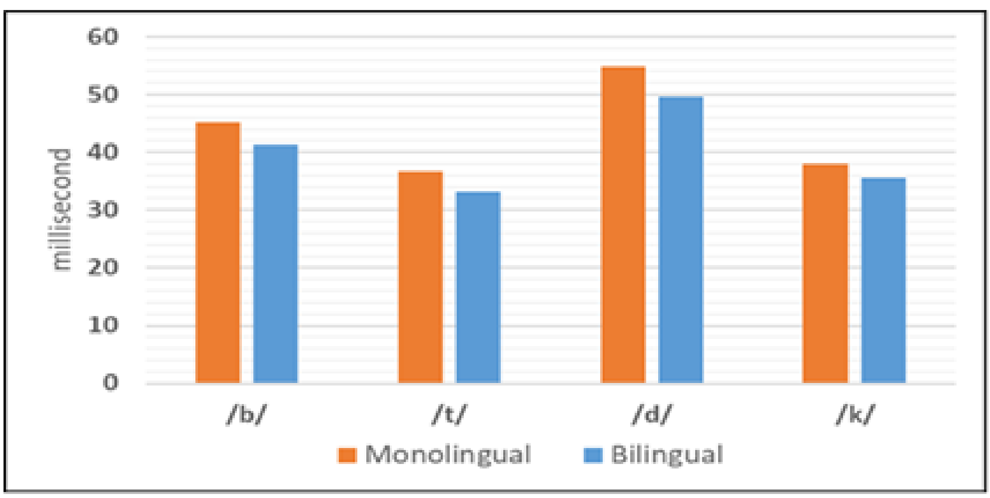

FIGURE 3. Mean PVD values in milliseconds of YA word-medial stops produced by monolingual and bilingual subjects

TABLE 4. Descriptive and Inferential Statistics for PVD of word-medial stops produced by monolingual and bilingual subjects

Sound	Monolingual		Bilingual		df	T	Sig
	Mean	Std. Deviation	Mean	Std. Deviation			
/b/	45.25	11.63	41.39	11.16	58	1.31	p =.195
/t/	36.63	11.28	33.14	9.94	58	1.27	p =.209
/d/	54.90	11.39	49.66	11.61	55	1.72	p =.091
/k/	37.96	11.04	35.54	9.71	58	.90	p =.371

Based on this output, it could be concluded that there were no significant differences in mean PVD values for YA stops /b, t, d, k/ in word-medial position produced by monolingual and bilingual subjects.

DISCUSSION

Acoustic analysis of YA stops shows that VOT is a sufficient cue in differentiating between voiced and voiceless stops. In initial and medial positions, both bilinguals and monolinguals had lag VOT values (positive values) for /t, k/ and lead VOT values (negative values) for /b, d/. This indicates that Yemeni Arabic follows the voicing lead and voicing lag pattern as in several other dialects of Arabic (Adam, 2012; Al-Nuzaili, 1993; Khattab, 2002; Kulikov, 2018; Rahim & Kasim, 2009; Rifaat, 2003; Tamim & Hamann, 2021; Yeni-Komshian, Grace H. Caramazza & Preston, 1977). Bilinguals showed mean VOTs for /b, d/ that are very close in values to the monolinguals. This indicates that in the production of /b, d/, there was no influence of L2 English on L1 Arabic in bilinguals.

On the contrary, for stops in both initial and medial position, bilinguals showed longer mean VOT values than monolinguals for the voiceless stops /t, k/. The differences in mean VOT values of /t, k/ in the initial position between bilinguals and monolinguals were significant. However, in the medial position, only /k/ showed significant differences. These higher mean VOT values by bilinguals indicate that they produced VOT values of /t, k/ that are in between monolinguals of both languages. Findings reveal that bilinguals showed signs of English-induced L1 phonetic drift in their pronunciation of Arabic /t, k/ whereas not in their realization of Arabic /b, d/. These results suggest equivalence classification between Yemeni Arabic and English for voiceless stops, but not voiced stops. Besides, mean VOT values of /t, k/ by monolinguals and bilinguals in this study were much higher in the word-initial position than in the word-medial position. This pattern was also found in Antoniou et al. (2010) where the acoustic measurement for a sound varied due to its position in a word.

Through previous studies on VOT, it has often been observed that two-category VOT languages are frequently reported to fall into the two adjacent zones along the VOT dimension: short lag - long lag or voicing lead - short lag but not voicing lead - long lag (as seen in this study). However, some exceptions have been reported whereby speakers produced fairly long "intermediate" lag values rather than short lag values (Abramson & Whalen, 2017; Flege & Port, 1981). However, in the foundational article by Lisker & Abramson (1964), there was no reference to this intermediate voicing lag. In this study, YA monolingual speakers showed mean VOT values for initial /t, k/ which fall within this intermediate zone whereas bilingual YA voiceless stops in initial position consistently showed mean VOT values in the long lag range. This finding of monolingual voiceless stops is consistent with results of Yemeni Arabic by Al-Nuzaili (2019) in which mean VOT values were in the intermediate lag range, as opposed to all the other Arabic dialects discussed earlier.

Furthermore, bilinguals showed shorter mean PVD values than monolinguals for all the stops involved. However, the differences were not statistically significant between the two groups. This indicates that L2 English acquisition did not cause any significant L1 phonetic drift in the duration of the preceding vowels investigated. Besides, voiced /b, d/ had longer mean PVD than voiceless /t, k/ in both groups.

CONCLUSION

The findings of this research reveal that Yemeni Arabic-English bilinguals demonstrate L1 phonetic drift towards their L2 English in the production of the voiceless stops /t, k/ but not in the production of voiced stops /b, d/. Such findings support the claim that phonetic properties of similar L2 categories can influence L1 phonological categories. The occurrence of L1 phonetic drift in adult L1 speakers proves that the L1 system remains dynamic and instable over the life span.

The fact that in this study, L1 phonetic drift was present in the VOT production of voiceless stops but not voiced stops might be due to Arabic voiceless stops and English voiced stops being in the lag region causing bilinguals to produce voiceless stops that have higher VOTs. This finding is interpreted as an attempt to retain cross-linguistic phonetic contrast between L1 Arabic voiceless stops and L2 English voiced stops within a bilinguals' common phonological space. On the contrary, Schwartz (2020) showed that L1 phonetic drift was evident in voiced stops but not voiceless stops of Polish- English bilinguals. Besides, Tobin et al. (2017) interpreted that not detecting any L1 drift in Spanish speakers learning L2 English after a couple of months of L2 immersion in the United States is due to a decline in L1 use.

Although earlier research has demonstrated that a reduction in L1 usage may contribute to phonetic drift, this was not the case in this study as the subjects were L1 dominant. Rather, L2 proficiency seemed to contribute to the drift in L1 production. However, for phonetic drift to occur, the amount of L2 proficiency required remains undetermined. Therefore, future studies need to shed light on the minimum amount of L2 English learning that is required to trigger a phonetic effect on L1 Yemeni Arabic. Besides, this study was limited to investigating proficient Arabic – English bilinguals. Since in Yemen, individuals are usually first exposed to English in classroom settings, future studies can focus on novice learners of English who are learning their L2 in formal classroom settings. Moreover, further research can investigate significant adjustments to L1 Yemeni Arabic in multilingual subjects. In other words, to what extent can a newly acquired L3 affect both L1 and L2. In addition, to control the dialectal effect on L1, this study was limited to the Taizzi dialect of Yemeni Arabic. Future research may investigate the acoustic properties of other dialects of Yemeni Arabic.

Furthermore, studies investigating bilingual speech have included an examination of the perception and/or production of both languages of a bilingual. However, this study investigated the production of YA with no reference to perception. Since accurate perception isn't always a requirement for accurate production (Shahidi et al., 2012), it is recommended that further research investigates the perception of Yemeni Arabic-English bilinguals.

REFERENCES

Abramson, A. S., & Whalen, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, *63*, 75–86. https://doi.org/10.1016/j.wocn.2017.05.002

Adam, H. (2012). VOT-Analysis: The Production of Stops by Agrammatic Palestinians. *International Journal of Linguistics*, *4*(4), 300–308. https://doi.org/10.5296/ijl.v4i4.2736

Al-Nuzaili, A. M. (1993). *Experimental study of emphasis and voicing in the plosives of Yemeni Spoken Arabic with some implications for foreign language teaching and learning*. The University of Leeds.

AlDahri, S. S. (2012). A study of voice onset time for modern standard Arabic and classical Arabic. *IEEE International Conference on Signal Processing, Communication and Computing*, 691–695. https://doi.org/doi:10.1109/icspcc.2012.6335582

AlDahri, S. S., & Alotaibi, Y. A. (2010). A crosslanguage survey of VOT values for stops (/d/, /t/). *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems*, *3*, 334–338. https://doi.org/10.1109/ICICISYS.2010.5658744

Alsiraih, W. (2020). Acoustic Analysis of Iraqi Arabic Stop Consonants. *Journal of Basra Research for Human Sciences*, *45*(2), 2–37.

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics*, *38*(4), 640–653. https://doi.org/10.1016/j.wocn.2010.09.005

Bergmann, C., Nota, A., Sprenger, S. A., & Schmid, M. S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics*, *58*, 71–86. https://doi.org/10.1016/j.wocn.2016.07.001

Boersma, P., & Weenink, D. (2019). *PRAAT: Doing Phonetics by Computer (Version 6.0.46). Computer Program*. (6.0.46).

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*(2), 249–268. https://doi.org/10.1016/j.wocn.2011.10.007

Dmitrieva, O., Jongman, A., & Sereno, J. A. (2020). The effect of instructed second language learning on the acoustic properties of first language speech. *Languages*, *5*(4), 1–33. https://doi.org/10.3390/languages5040044

Docherty, G. J. (1992). *The Timing of Voicing in British English Obstruents*. Walter de Gruyter.

Flege, J. E. (1995). Second Language Speech Learning: Theory, Findings, and Problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 233–277). York Press.

Flege, J. E., & Port, R. (1981). Cross-Language Phonetic Interference: Arabic to English. *Language and Speech*, *24*(2), 125–146. https://doi.org/http://dx.doi.org/10.1177/002383098102400202

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*(4), 467–491. https://doi.org/10.1016/S0167-6393(02)00128-0

Fox, R. A., & Jacewicz, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. *The Journal of the Acoustical Society of America*, *126*(5), 2603–2618. https://doi.org/10.1121/1.3212921

Harada, T. (2003). L2 Influence on L1 Speech in the Production of VOT. In and J. R. M.-J. Sol´e, D. Recasens (Ed.), *Proceedings of the 15th International Congress of Phonetic Sciences*

(pp. 1085–1088). Causal Productions.
Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2016). Mutual influences between native and non-native vowels in production: Evidence from short-term visual articulatory feedback training. *Journal of Phonetics*, *57*, 21–39. https://doi.org/10.1016/j.wocn.2016.05.001
Khattab, G. (2002). VOT Production in English and Arabic Bilingual and Monolingual Children. In D. B. Parkinson & E. Benmamoun (Eds.), *Perspectives on Arabic Linguistics XIII-XIV* (pp. 1–37). John Benjamins Publishing Company.
Kulikov, V. (2018). *Laryngeal contrast in Qatari Arabic : Effect of speaking rate on VOT*. https://doi.org/10.13140/RG.2.2.18212.17284/2
Ladefoged, P. (2003). *Phonetic data analysis*. Blackwell Publishing.
Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD*, *20*(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830
Mackay, I. R. A., Meador, D., & Flege, J. E. (2001). The Identification of English Consonants by Native Speakers of Italian. *Phonetica*, *58*, 103–125. https://doi.org/10.1159/000028490
Mayr, R., Price, S., & Mennen, I. (2012). First language attrition in the speech of Dutch-English bilinguals: The case of monozygotic twin sisters. *Bilingualism*, *15*(4), 687–700. https://doi.org/10.1017/S136672891100071X
Mayr, R., Sánchez, D., & Mennen, I. (2020). Does teaching your native language abroad increase L1 attrition of speech? The case of spaniards in the United Kingdom. *Languages*, *5*(4), 1–20. https://doi.org/10.3390/languages5040041
Mitleb, F. M. (2009). Voice Onset Time of Jordanian Arabic Stops. *3rd International Conference on Arabic Language Processing (CITALA'09), May 4-5, 2009, Rabat, Morocco*, 133–135.
Mora, J. C., Keidel, J. L., & Flege, J. E. (2015). Effects of L2 use on the production of L1 vowels in early Spanish-Catalan bilinguals. In M. Romero, J. & Riera (Ed.), *The Phonetics-Phonology Interface. Representations and methodologies* (pp. 33–53). John Benjamins Publishing Company.
Naji, R. A. M. (2019). An Acoustic Measurements of English Stops Produced by Native Yemeni Arabic Speakers of English (NYASE). *International Journal of Language and Literary Studies*, *1*(2), 88–109. https://doi.org/10.5281/ZENODO.2875276
Osborne, D. M., & Simonet, M. (2021). Foreign-language phonetic development leads to first-language phonetic drift: Plosive consonants in native portuguese speakers learning English as a foreign language in Brazil. *Languages*, *6*(3). https://doi.org/10.3390/languages6030112
Rahim, A. J., & Kasim, Z. R. (2009). A Spectrographic Study of Voice Onset Time in Arabic. *Journal of Education and Science*, *2*(16), 28–41.
Rifaat, K. (2003). Voice Onset Time in Egyptian Arabic: A Case where Phonological Categories Dominate. *Proceedings of the 15th International Congress of Phonetic Sciences*, 791–794.
Salem, N. M., & Pillai, S. (2020). An Acoustic Analysis of Intonation in the Taizzi variety of Yemeni Arabic. *Linguistics Journal*, *14*(1), 154–183.
Schwartz, G. (2020). Asymmetrical cross-language phonetic interaction. *Linguistic Approaches to Bilingualism*, *October*, 1–30. https://doi.org/10.1075/lab.19092.sch
Shahidi, A. H., & Aman, R. (2011). An acoustical study of English plosives in word initial position produced by Malays. *3L: Language, Linguistics, Literature*, *17*(2), 23–33.
Shahidi, A. H., Aman, R., & Kechot, A. S. (2012). Production and perception of English word

final stops by Malay speakers. *GEMA Online Journal of Language Studies*, *12*(4), 1109–1125.

Tamim, N., & Hamann, S. (2021). Voicing contrasts in the singleton stops of Palestinian Arabic: Production and perception. *Proceedings of the Annual Conference of the International Speech Communication Association*, *INTERSPEECH*, *6*, 401–405. https://doi.org/10.21437/Interspeech.2021-1079

Tobin, S. J., Nam, H., & Fowler, C. A. (2017). Phonetic drift in Spanish-English bilinguals : Experiment and a self-organizing model. *Journal of Phonetics*, *65*, 45–59. https://doi.org/10.1016/j.wocn.2017.05.006

Yeni-Komshian, Grace H. Caramazza, A., & Preston, M. S. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics*, *5*, 35–48. https://doi.org/10.1016/S0095-4470(19)31112-X

ABOUT THE AUTHORS

Sumaya Faisal Alshamiri is a postgraduate student at the Center for Research in Language and Linguistics, Faculty of Social Science and Humanities, Universiti Kebangsaan Malaysia. She is currently doing her doctorate degree in Experimental Phonetics.

Shahidi A.H. (Ph.D) is an Assoc. Prof. at the Center for Research in Malay Language, Literature and Culture, Faculty of Social Science and Humanities, Universiti Kebangsaan Malaysia. His research interests are related to Experimental Phonetics, Second Language Learning and Malay Dialects.

Sharifah Raihan Syed Jaafar (Ph.D) is a senior lecturer at the Center for Research in Language and Linguistics, Faculty of Social Science and Humanities, Universiti Kebangsaan Malaysia. Her research interests include Theoretical Phonology and Minority Language.