

An acoustic analysis of the production of word-initial stop /p/ by late Arab bilinguals

by Sawsan Askoul

Sawsan Askoul University of London sawsan.askoul@kcl.ac.uk

Published in *Training, Language and Culture* Vol 1 Issue 4 (2017) pp. 68-83 doi: [10.29366/2017tlc.1.4.5](https://doi.org/10.29366/2017tlc.1.4.5)

Recommended citation format: Askoul, S. (2017). An acoustic analysis of the production of word-initial stop /p/ by late Arab bilinguals. *Training, Language and Culture*, 1(4), 68-83. doi: [10.29366/2017tlc.1.4.5](https://doi.org/10.29366/2017tlc.1.4.5)

The aim of this study is to examine the production of the word-initial stop /p/ by Arabs speaking English as a second language (L2). Arabic was chosen as a counter language since its phonetic contrasts between voiced and voiceless sounds differ from English. The data was collected from 50 participants, consisting of one group of highly qualified Arabs (n= 30 bilinguals) living in London for 10 years, and two groups of baseline speakers (n= 10 for inexperienced native Arabic and native English baseline groups, respectively). The production of /p/ by late Arab bilinguals (who learned English after puberty) was elicited via three oral tasks. Acoustic analysis was conducted to measure the voice onset time (VOT) value of /p/. VOT is defined as the length of time between the release of a stop consonant and the onset of vocal fold vibration. Research shows that length of residence (LOR), level of education, quality and quantity of input and rigorous usage of L2 are more important than AOA (Age of Acquisition) in enabling L2 learners to achieve native speaker-like pronunciation of the sound /p/, although for late Arab bilinguals the overall results demonstrated that AOA is a key predictor of L2 sounds. The final results demonstrate that the production of /p/ by late Arab bilinguals differed significantly from that produced by inexperienced Arabs. Late Arab bilinguals managed to establish new phonetic categories for English /p/ that does not exist in their language (L1). However, their production of /p/ deviated from that of native English speakers. Only nine out of 30 late Arab bilinguals were able to achieve native speaker-like pronunciation of /p/.

KEYWORDS: pronunciation, bilingualism, late learners, Arabs, voiceless consonant, voice onset time, (VOT)



This is an open access article distributed under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0)

1. INTRODUCTION

In second language (L2) speech the production of some L2 sounds is challenging to L2 learners, especially those sounds that exist in the L2 speech system but do not exist in their first language (L1) inventory (e.g. Lado, 1957; cf. Flege, 1995; Flege & Port, 1981). The Age of Acquisition (AOA), defined as the age of the first exposure to an L2 (Flege et al., 1999), is the most studied variable

which is believed to affect L2 speech production. It has also been confirmed by a substantial amount of research as a key predictor of successful L2 attainment. The younger a learner is exposed to an L2, the more native speaker-like proficiency they attain (e.g. Abrahamsson & Hyltenstam, 2009; Baker, 2010; Bialystok, 1997; Birdsong, 2006; Bongaerts et al., 1997; DeKeyser, 2000; Johnson & Newport, 1989; Flege, 1991; Flege et al., 1995a;

‘The Age of Acquisition (AOA), defined as the age of the first exposure to an L2 (Flege et al., 1999), is the most studied variable which is believed to affect L2 speech production’

Flege et al., 1999; Muñoz & Singleton, 2011; Patkowski, 1990; Saito, 2015b; Scovel, 1988).

However, other studies speculate that successful L2 pronunciation is dependent on a wide variety of variables other than AOA such as L1 influence (Flege & Davidian, 1984), the distance between L1 and L2 (Flege, 2003), the relative usage of L1 and L2 (Flege, 1999; Flege & MacKay, 2004), the quality and quantity of L2 input (Flege & Liu, 2001; Jia & Aaronson, 2003), aptitude (Abrahamsson & Hyltenstam, 2008; Loup et al., 1994), motivation (Mayor, 1999), ethnic identity (Gatbonton et al., 2011; Gatbonton & Trofimovich, 2008), level of education (Derwing & Munro, 2008; Flege et al., 1999) and length of residence (LOR) in an L2 speaking community (e.g. Saito & Brajot, 2013; Saito, 2015a).

Nonetheless, the demonstrable variations shown by late learners (those who commenced L2 learning after puberty) in successfully mastering L2 sounds have suggested that the independent and/or combined effects of AOA, together with other

variables on L2 sound production, are largely inconclusive.

2. MATERIALS AND METHODS

Seeking to contribute to the growing body of research in the area of L2 speech acquisition, this study was conducted with the aim of examining the production of the English /p/ by Arabic speakers who have learned English from the age of 13 onwards. It was accomplished by performing an acoustic analysis of voice onset time (VOT) – defined as the length of time between the release of a stop consonant and the onset of the vocal fold vibration (Lisker & Abramson, 1964, 1967) in their speech, alongside other acoustic cues which are not discussed in this article. Our aim was to determine whether late Arab bilinguals are able to produce the sound /p/, which does not exist in their L1 inventory, like a native speaker VOT value. The sound of /p/ produced by late Arab bilinguals was then compared to that of native speakers. The tendency is for the native speaker /p/ sound to be replaced by a /b/ sound by late Arab bilinguals. A bachelor’s degree obtained from an English university was set as the minimum qualification required in order to participate in the study, in addition to, LOR (≥ 10 years), intensive and authentic input from English native speakers (average 70%) and intensive English use on a daily basis (average 85%). We also hypothesise that AOA is not a key predictor of L2 sound production by late Arab participants. Therefore, it

may safely be neglected when other variables have been firmly controlled, specifically, level of education, LOR, quality and quantity of L2 input and dominant usage of L2. Such variables combined allow late Arab bilinguals to attain native speaker-like pronunciation of the sound /p/ with a native speaker-like VOT value, despite their late AOA (≥ 13 years).

3. STUDY AND RESULTS

3.1 Voice Onset Time (VOT)

There is broad consensus that foreign accents are a natural outcome of late L2 learning. This notion is supported by research and empirical studies demonstrating that late learners are most likely to produce their L2 speech with non-native patterns of pronunciation (Derwing & Munro, 2005). However, native speakers of a language are equipped with the ability to discriminate whether or not a speaker's accent reflects their mother tongue, even when the speech is structured with faultless syntax and a robust lexicon. This judgement is frequently made by detecting certain acoustic cues and phonetic features produced by the speaker (Scovel, 1995). Although there are different acoustic features that distinguish speech sounds, such as articulatory force (*fortis* and *lenis*) and aspiration (when a sound is pronounced with/ out an audible explosion and an interlude of noise), voicing is undoubtedly one of the most salient auditory signals '*identifiable by ear with great ease*' (Lisker & Abramson, 1967).

'However, native speakers of a language are equipped with the ability to discriminate whether or not a speaker's accent reflects their mother tongue, even when the speech is structured with faultless syntax and a robust lexicon'

In contrast, articulatory force has no agreed physical meaning and aspiration has a contrastive meaning limited to particular contexts. Voicing is readily measurable and a distinction between the contrasting sounds in all positions of occurrence is possible (Lisker & Abramson, 1967; cf. Yeni-Komshian et al., 1977). Voicing distinguishes between voiced sounds that are characterised by vocal cord vibration or glottal pulsing (Lisker & Abramson, 1964, 1967) and voiceless sounds that are pronounced without such glottal pulsing. The occurrence or non-occurrence of pulsing is a distinctive feature that is more often accompanied by stop consonants than is the case for other speech sounds (Yu et al., 2015). English stop consonants include the voiced /b/, /d/ and /g/ that cognate to the voiceless /p/, /t/ and /k/. Such sounds are pronounced by closure or severe constriction and the build-up of considerable air pressure in the oral cavity, followed by a sudden release of the constriction (Flege, 1988; Lisker & Abramson, 1964, 1967).

‘Voicing is readily measurable and a distinction between the contrasting sounds in all positions of occurrence is possible’

When acoustically analysed, the closure of stop consonants is fairly well marked and clearly shown on spectrograms by the presence of a ‘burst’ or brief interval of high-intensity noise (Lisker & Abramson, 1967). The onset of the burst can be fixed with considerable certainty and provides a convenient reference point for measuring voicing duration. Consequently, it is possible to relate the degree of voicing of a stop to the time between the burst and onset of pulsing. This time is known ‘voice onset time’ (VOT). VOT is defined as the duration of the period of time between the beginning of the stop release burst and the onset of the vocal fold vibration, glottal pulsing or periodicity, as some researchers prefer to call it. It is usually measured in milliseconds (ms). VOT is commonly found in different languages around the world and has three universal categories: voicing lead (50-150ms), short lag (0-30ms), and long lag (50-110ms).

3.2 English /p/

In terms of phonology, the sound /p/ is classified as one of the English phonemes which, according to the *Longman Dictionary of Linguistics*, is the

smallest unit of sound in a language which can distinguish two words. In English, the words *pan* and *ban* differ only in their initial sounds: *pan* begins with /p/ and *ban* with /b/, therefore, /p/, /b/ are phonemes. Since phonemes are categories rather than actual sounds, they are not tangible, instead they are abstract theoretical groups that have only a psychological reality. In other words, we cannot hear phonemes, but we assume they exist because of the way they are used by native speakers (e.g. Flege & Port, 1981; Murray, 1995).

In comparing English to Arabic sound systems, Flege (1981) has observed that /p/ is considered ‘voiceless’ because of the absence of glottal pulsing upon articulation. No equivalent sound to /p/ is found in the Arabic sound system (Yeni-Komshian et al., 1977), which makes it an odd language among its sister Semitic languages (Newman, 2002).

The closest Arabic sound to the English /p/ is [پ]. [پ] is pronounced with vibration of the vocal folds and a very short obstruction of the flow of air as the two lips are pressed together but the velum is up, thus blocking access to the nasal cavity. The pressure builds up behind the blockage caused by the two lips and the lips come apart and the air escapes through the mouth in a burst (Newman, 2002).

In conclusion, Arabic lacks an equivalent to the

‘In terms of phonology, the sound /p/ is classified as one of the English phonemes which, according to the ‘Longman Dictionary of Linguistics’, is the smallest unit of sound in a language which can distinguish two words’

English /p/ which constitutes a big challenge for any Arabic speaker wishing to pronounce the English /p/ accurately, especially for those who learn English in naturalist settings, as the distinction requires a high degree of perception and awareness of details which differentiate the English /p/ from the Arabic [پ].

3.3 Participants

The participants were chosen on the basis of the results of an initial interview (Saito & Brajot, 2013). In order to control other variables that could affect the Arabic speakers’ oral production, the participants were required to meet the following criteria: a) native Arabs, b) highly qualified with a bachelor degree minimum obtained from England, c) engaged in full time jobs, d) living in England for a minimum of 10 years, e) first intensive exposure to English was at the age of 13 years or more, and f) bilinguals spoke two languages: Arabic as an L1 and English

as an L2. This last condition was with the aim of excluding participants from multilingual backgrounds.

3.4 Baseline speakers

Data was also collected from two control groups to determine baseline acoustic description for native Arabic and English speakers’ production of the English /p/.

The Arabic speaking baseline was determined by highly qualified Arabs (e.g. Flege, 1991) with educational attainments and work experience equivalent to those of late Arab bilinguals. In their country of origin, Syria, the Arabic speaking baseline participants worked in: a) educational sectors – one headteacher, teachers and two retired teachers, b) engineering – agricultural and civil engineers, and c) university students. There were 10 native speakers of Arabic (7 females and 3 males) who had just arrived in London with LOR ranging from 3 to 8 months only. Their ages ranged from 19 years to 70 years (mean age 40.4 years, SD=20.08). As can be seen, the chronological age of the Arabic speaking baseline participants was closer to the age of the late bilinguals than those in the English baseline (see below).

Regarding the English baseline and following Abrahamsson & Hyltenstam’s (2009) model in selecting baseline groups, there were 10 university students, a) some speaking English at home during

childhood, b) with English as the only language at school, c) one who had lived his whole life in England without interval. In order to ensure that a variety of VOT values differing according to age, sex and speech rate (Yu et al., 2015) was consistent with the variety reflected in the Arabic speaking baseline, the English baseline consisted of seven females and three males. Additionally, in an attempt to match the Arabic speaking baseline and late bilinguals' chronological age to that of the English baseline participants, one elderly retired teacher participated to constitute the group. Their ages ranged from 19 -70 years old (mean age 30, SD=14.8). Linguistically they had no previous knowledge of English phonetics and phonology.

3.5 Target words

The 12 target words exploited in the three oral tasks had the form of consonant-vowel-consonant (CVC) word-initial (e.g. Flege & Eefting, 1987b; Flege & Liu, 2001; Flege et al., 1995a) and singletons (e.g. Flege & Brown, 1982). The lexical familiarity and frequency (Flege et al., 1996) of the words were taken into consideration on the grounds that lexical factors are believed to have some effects on L2 sound production (Saito & Brajot, 2013).

3.6 Task description

Echoing Saito & Brajot (2013), the selection of tasks utilised in this study aimed to focus on late Arab bilinguals' /p/ production in more

'Arabic lacks an equivalent to the English /p/ which constitutes a big challenge for any Arabic speaker wishing to pronounce the English /p/ accurately'

communicative contexts within a '*realistic time limit*' (Saito & Brajot, 2013, p. 851). They were not allowed prolonged planning times, so that they would demonstrate their real implicit knowledge stored in long-term memory (Flege, 2015) rather than exhibiting explicit knowledge stored in general memory (Ellis, 2005). The three tasks included:

1. *A series of picture descriptions.* In order to elicit the natural spontaneous speech of the participants without causing them to pay attention to their production of /p/, the participants were presented with a cartoon used in Munro and Derwing's (1999) study that illustrated '*an amusing story*' (p. 290) and were asked to describe the events depicted. The cartoon was displayed on a classroom projector or a laptop screen with seven word prompts, three of which were target words and the other four of which were distractors (Saito & Brajot, 2013). The target words included; *pass*, *pain*, and *pick* while the distracting words were *city*, *suitcase*, *man* and *women*. Each participant was given 10 seconds of planning time for this kind of task.

2. *Timed picture descriptions.* In this task, the participants had to describe what they saw in seven different pictures. Under each picture, three prompt words were given. One of the three words was the target word. Overall, three out of the seven pictures contained the target words, and the other four pictures were distractors. For example, the distracting picture contained a picture of a cat sitting in a sink. The cat had noticeably big green eyes. The three prompt words were *cat*, *sink* and *eyes*. The pictures were in a certain order, whereby the distractors were presented first then the pictures with the target words. This technique was used with the purpose of familiarising the participants with the procedure (Saito & Brajot, 2013). They were not required to say more than two to three sentences to describe each picture using the target words which included *pill*, *pen* and *pan*. The pictures were presented on a classroom projector or a laptop screen. Participants were given five seconds of planning time after each picture had been displayed.

3. *Word reading.* In this task, the participants read a list of words (e.g. Yeni-Komshian et al., 1977; Riney & Takagi, 1999) consisting of 17 words. Overall, eight words were target words and nine other words functioned as distractors. The target words were *put*, *pass*, *pull*, *push*, *poor*, *pet*, *pain* and *pack*. The distractors included frequent and easy words such as *snow*, *window* and *mobile*, in addition to less frequent words such as *linguistics*,

‘English /p/ which constitutes a big challenge for any Arabic speaker wishing to pronounce the English /p/ accurately, especially for those who learn English in naturalist settings, as the distinction requires a high degree of perception and awareness of details which differentiate the English /p/ from the Arabic [ب]’

economics, *industry* and *geography* (Saito & Brajot, 2013). The word list was presented on a classroom projector or a laptop screen with a font size of 72. No planning time was given for this task.

3.7 Acoustic analysis

All 50 participants, late Arab bilinguals (n=30) and the 2 baseline groups (n=10 in each group) completed the three tasks in the same order, in addition to an initial interview prior to the tasks. This took place at a convenient location for each participant, including their workplace, home and university classroom. The instructions for all procedures were delivered in English for the English baseline group, and Arabic for both late Arab bilingual and Arabic speaking baseline groups (cf. Flege, 1980; Flege & Eefting, 1987a). The speech production by all participants was

'The instructions for all procedures were delivered in English for the English baseline group, and Arabic for both late Arab bilingual and Arabic speaking baseline groups'

recorded either directly on Praats (the acoustic analyser of speech sounds downloaded from its official website with its updated version (last modified in May 2017), set at 48000 Hz, which is appropriate for human speech on a portable laptop with a unidirectional microphone) or by using an Olympus WS 812 4GB voice recorder. Audio clips were saved on the same laptop in waveform audio. Echoing Saito and Brajot (2013), in order to obtain spontaneous speech production and to prevent them from focusing on their pronunciation, the participants were informed that the oral tests were to check the fluency and comprehensibility of their speech. However, the real aim of the study was unveiled once they had finished the tasks. The tests were given in the order of a series of pictures description, timed picture description and finally word reading. Following Flege et al.'s (1995a) model, the VOT of the sound /p/ in all tokens (12 initial-stop words x 30 Arab late bilinguals x 10 English baseline x 10 Arabic speaking baseline) was measured spectrographically using a Praat spectrogram and waveform.

4. DISCUSSION

Following Stölten et al.'s (2014) model, the mean values of VOT for all words produced in the three oral tests (a series of pictures description, timed picture description and word reading tests) were submitted to one-way ANOVA to establish if there was any significant difference between the three groups (late Arab bilinguals, Arabic speaking and English baselines). ANOVA is a statistical variance analysis test developed by Ronald Fisher in 1918. The ANOVA results indicated that there were overall significant differences between the three groups ($F(2, 47) = 52.3, p = 3.19$). However, the one-way ANOVA test usually shows there is an overall difference between groups, but it does not state which specific group differs. For that reason, an additional confidence interval (CI) test was also carried out on the mean value of VOT produced by each group. The results for the Arabic speaking baseline, late Arab bilinguals and English baseline, (95% CIs [7.9, 10.1], [49.4, 62.8] and [66.2, 82.8]) respectively, demonstrated that the VOT values of /p/ produced by late Arab bilinguals differed considerably from the VOT values produced by the Arabic speaking baseline, which was expected. The VOT values of /p/ by late Arab bilinguals also deviated significantly from that of the English baseline, which was unpredicted. Furthermore, no overlap was observed between the English baseline and late Arab bilinguals' VOT values as was further demonstrated by a post hoc t-test ($p < 0.001$) (see Figure1).

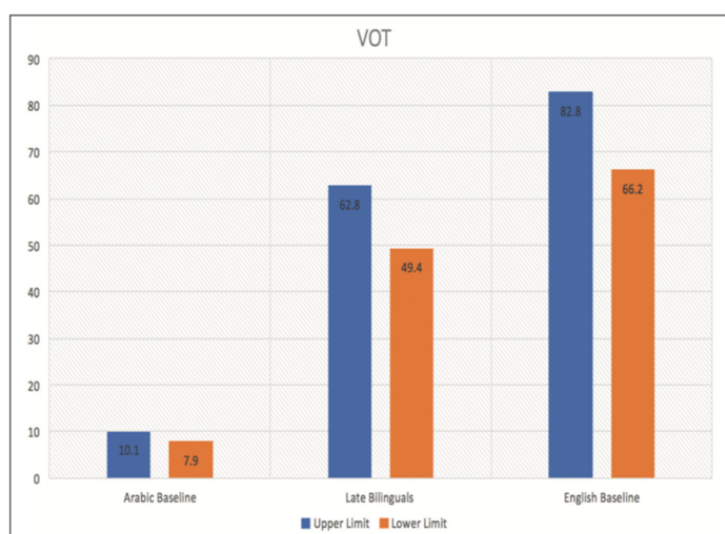


Figure 1. The result of CIs from the late Arab bilinguals group compared with the results of baseline groups

Saito and Brajot (2013) emphasise that L2 learners perform better in word reading than spontaneous speech as they may have sufficient time to access their explicit knowledge, rather than tapping into their actual implicit knowledge, resulting in production of more native speaker-like L2 sounds. On the other hand, Lisker and Abramson (1967) observe that the VOT value of stop consonants in all languages changes depending on whether /p/ occurs in running speech or isolated words (also see Kessinger & Blumstein, 1998). In order to confirm Lisker and Abramson's (1967) observation, and echoing Yeni-Komshian et al., (1977), three different correlated t-tests were conducted to compare VOT values between spontaneous speech tasks and word reading tasks within each group. The results confirmed Lisker and Abramson's (1967) observation but it did not demonstrate how

well or poorly late Arab bilinguals did in comparison to the two baselines. To achieve this insight and following Flege's (1991) model, two one-way ANOVA tests were carried out to compare the performance of late Arab bilinguals to those of the Arabic speaking and English baselines in the spontaneous speech tasks. Another two one-way ANOVA tests were run to compare the performance of late Arab bilinguals to those of the Arabic speaking and English baselines in the word reading task. The results demonstrated that late Arab bilinguals' performance was much better than that of the Arabic speaking baseline in word reading ($F(1,38) = 52.5, P = 1.13$). However, it also differed from the performance of the Arabic speaking baseline in spontaneous speech ($F(1,38) = 48.1, P = 2.9$). The overall results established that late Arab bilinguals delivered a much better

performance than did the Arabic speaking baseline in the three tasks, as hypothesised. Yet, they were unable to completely match the English baseline in both spontaneous speech and word reading tasks.

Computing the mean VOT value of each participant revealed that nine participants produced mean VOT values similar to the VOT values of the English baseline. A subgroup for those participants was created and termed 'Highly Professional'. Following the same procedure as that conducted for the other participants, the results from a one-way ANOVA indicated that the

mean VOT value of the 'highly professional' subgroup and the English baseline were almost equal.

A further CI test was conducted to accurately identify the upper limits and the lower limits of VOT values in each group. The results indicated that the upper limits of the two groups did not differ greatly. However, the lower limit of the highly professional group was higher than that of the English baseline (see Figure 2), which means there was a slight overlap between the two groups. This overlap was not greatly significant thus both groups produced /p/ in a very similar manner.

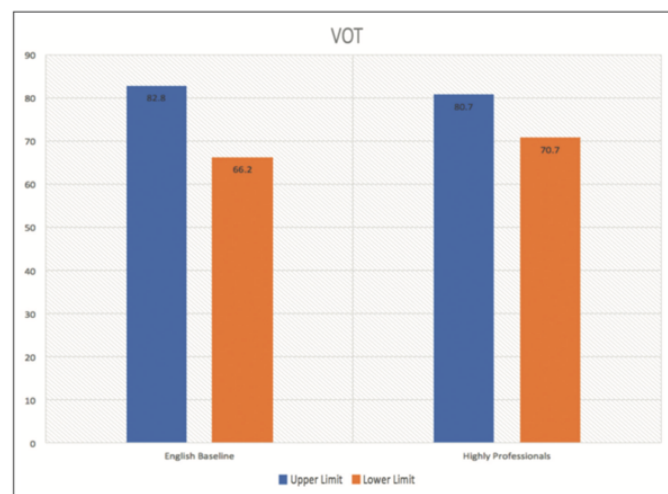


Figure 2. The results of CIs from the highly professional group (n=9) compared with the results of English baseline

5. CONCLUSION

The study measured VOT initial /p/ sounds produced by late Arab bilinguals through three different oral tasks. This was accomplished to prove

the hypothesis that AOA is not a key predictor when other variables are firmly controlled, namely LOR, level of education, quality and quantity of L2 input and rigorous usage of L2. The overall results

‘Although there has been a shift from a narrow focus on linguistic competence to a broader focus on communicative competence, teaching pronunciation is still underestimated and has a marginalised status compared to teaching other skills’

demonstrated that AOA is a key factor in predicting late Arab bilinguals’ proficiency. Brain maturity was judged to be the key factor in predicting the overall quality of late Arab bilinguals’ /p/ production. Indeed, as Bley-Vroman (1989) observe, *‘Virtually no adult learner achieves perfect success, if what one means thereby is development of native speaker competence’* (Bley-Vroman, 1989, p. 44).

Despite the common observation that L2 learners usually strive to attain native-speaker linguistic ability, few adult learners can attain this goal, and a distinct foreign accent is a common characteristic of L2 speech (Abrahamsson & Hyltenstam, 2009; Flege et al., 1995a). Foreign accents are usually described by omissions, substitutions, or distortions of L2 sounds (Flege, 1988). One instance of such distortions is the distortion of the VOT value of stop consonants.

Empirical studies have demonstrated that L2 learners’ production of VOT is closely correlated with native speakers’ perceptions of foreign accent. Major (1987) measured VOT values by adult learners learning English as a foreign language in Brazil. He found a positive correlation between VOT value and foreign accent; the more native speaker-like a speaker is rated, the closer the VOT conforms to the English norms. Similar results have been achieved by other researchers such as Flege and Eefting (1987b) and Riney and Takagi (1999). Accordingly, it is firmly established that ‘in L2 pronunciation there is a basic correlation between GFA (global foreign accent) and VOT’ (Riney & Takagi, 1999, p. 298).

Although many researchers argue that L2 learners should set up a realistic goal for prioritising comprehensibility over attaining native speaker-like proficiency (e.g. Abrahamsson & Hyltenstam, 2009; Derwing & Munro, 2008; Flege et al., 1995a), the case of Arab learners may be different, as this study shows. Many experimental studies confirm that when Arabs speak English, they tend to produce /p/ with a very small VOT value resembling that of /b/. In a study by Flege and Port (1981), English native judges perceived /p/ produced by Arabs as /b/, since Arab participants produced the voiceless /p/ as a voiced /b/ with a short VOT value that did not exceed 30ms. The Arab participants in that study produced the word *bin* instead of *pin*, for example, when they were

‘Accordingly, it is firmly established that in L2 pronunciation there is a basic correlation between GFA (global foreign accent) and VOT’

instructed to produce /p/ in initial word positions. Therefore, such instances are not only about the substitution of one sound with another, rather they involve replacing one word with another that has a different meaning. This could affect their ability to produce comprehensible output and deter effective communication causing listeners to misjudge Arabs’ affective state or provoking negative personal evaluation. This is further proved by Arabs’ self-assessment and anecdotes.

Although there has been a shift from a narrow focus on linguistic competence to a broader focus on communicative competence (Morley, 1991), teaching pronunciation is still underestimated and has a marginalised status compared to teaching other skills (Derwing & Munro, 2005).

In her helpful article, Morley (1991) provides a long list of advice and recommendations to educational institutions, teachers and learners

equally. Among the different recommendations, she focuses on the necessity to equip teachers, through the necessary training, with a very specific knowledge of applied English phonetics and phonology, in addition to increasing the number of professionals with informed expertise directed toward facilitating learners’ development of functional pronunciation patterns. She also asserts the continuing need for further research into aspects of language sounds with a distinguishable phonological and phonetics nature, as well as the course of development of L2 phonological system and different inter-language phonologies. She discusses the technology revolution that has been very advantageous to pronunciation work, in supplying teachers with a variety of teaching materials such as videos along with computer capacity applicable in classroom settings and learning laboratories.

Additionally, she states that learners themselves should take certain responsibility for improving their pronunciation through self-awareness of the features of speech production and speech performance and by building a personal repertoire of speech monitoring and modification skills in order to continue to improve speaking effectiveness in English in the long run, among many other recommendations.

Editor’s note: For a longer version with additional research, please apply to us.

References

- Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, 30(4), 481-509.
- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(2), 249-306.
- Abramson, A., & Lisker, L. (1965). *Voice onset time in stop consonants: Acoustic analysis and synthesis*. New York, NY: Haskins Laboratories.
- Baker, W. (2010). Effects of age and experience on the production of English word-final stops by Korean speakers. *Bilingualism: Language and Cognition*, 13(3), 263-278.
- Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research*, 13(2), 116-137.
- Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56(1), 9-49.
- Birdsong, D. (2014). Dominance and age in bilingualism. *Applied Linguistics*, 35(4), 374-392.
- Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. Gass, & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 41-68). New York, NY: Cambridge University Press.
- Boersma, P., & Weenink, D. (2011). *Praat: Doing phonetics by computer*. Retrieved from <http://www.praat.org>.
- Bohn, O., & Flege, J. (1993). Perceptual switching in Spanish/English bilinguals: Evidence for universal factors in stop voicing judgments. *Journal of Phonetics*, 21, 267-290.
- Bongaerts, T., Van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447-465.
- Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern-day 'Wild Child'*. New York, NY: Academic Press.
- Curtiss, S., Fromkin, F., Krashen, S., Rigler, D., & Rigler, M. (1974). The linguistic development of Genie. *Linguistic Society of America*, 50(3), 528-554.
- DeKeyser, R. M. (2000). The robustness of Critical Period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499-533.
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379-397.
- Derwing, T. M., & Munro, M. J. (2008). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476-490.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27(2), 141-172.
- Flege, J. (1980). Phonetic approximation in second language acquisition. *Language Learning*, 30(1), 117-134.
- Flege, J. (1987). The production of 'new' and 'similar' phones in a foreign language: Evidence for the Effect of Equivalence classification. *Journal of Phonetics*, 15, 47-65.
- Flege, J. (1991). Age of learning affects the authenticity of Voice Onset Time (VOT) in stop consonants produced in a second language. *Acoustic Society of America*, 89(1), 395-411.
- Flege, J. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America*, 93(3), 1589-1608.
- Flege, J. (1995). Second language speech learning: Theory, findings & problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.
- Flege, J. (1999). *Age of learning and second-language speech*. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period hypothesis* (pp. 101-132). Hillsdale, NJ: Lawrence Erlbaum.
- Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer, & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences*

- and similarities (pp. 319-355). Berlin, Germany: Mouton de Gruyter.
- Flege, J. (2005, 14-15 April). Origins and development of the Speech Learning Model. *Keynote lecture presented at the 1st ASA Workshop on L2 Speech Learning*. Vancouver, Canada: Simon Fraser University.
- Flege, J., & Brown, W. S. (1982). The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterances. *Journal of Phonetics*, 10, 335-345.
- Flege, J., & Davidian, R. (1984). Transfer and development processes in adult foreign language speech production. *Applied Psycholinguistics*, 5, 323-347.
- Flege, J., & Eefting, W. (1986). Linguistic development effects on the production and perception of stop consonants. *Phonetica*, 43, 155-171.
- Flege, J., & Eefting, W. (1987a). Cross-language switching in stop consonant perception and production by Dutch speaker of English. *Speech Communication*, 6, 185-202.
- Flege, J., & Eefting, W. (1987b). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, 15, 67-83.
- Flege, J., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Acoustical Society of America*, 83(2), 729-740.
- Flege, J., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23(4), 527-552.
- Flege, J., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26, 1-34.
- Flege, J., Moeini, S., & McCutcheon, M. J. (1988). The design of a microcomputer-controlled Voice Onset Time analyser. *Biomedical Instrumentation & Technology*, 357-362.
- Flege, J., Munro, M., & MacKay, I. R. A. (1995a). Effects of age of second-language learning on the production of English consonants. *Speech Communication*, 16(1), 1-26.
- Flege, J., Munro, M., & MacKay, I. R. A. (1995b). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97(5), 3125-3134.
- Flege, J., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech*, 24(2), 125-146.
- Flege, J., & Schmidt, A. M. (1995). Native speakers of Spanish show rate-dependent processing of English stop consonants. *Phonetica*, 52(2), 90-111.
- Flege, J., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *Journal of Acoustic Society of America*, 99(2), 1161-1173.
- Flege, J., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second language acquisition. *Journal of Memory & Language*, 41, 78-104.
- Fourakis, M., & Iverson, G. K. (1984). On the acquisition of second language timing patterns. *Language Learning*, 35(3), 431-442.
- Gass, S. (1996). Second language acquisition and linguistic theory: The role of language transfer. In W. Ritchie, & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 384-403). San Diego, CA: Academic Press.
- Gatbonton, E., & Trofimovich, P., & Segalowitz, N. (2001). Ethnic group affiliation and patterns of development of a phonological variable. *The Modern Language Journal*, 95(2), 188-204.
- Gatbonton, E., & Trofimovich, P. (2008). The ethnic group affiliation and L2 proficiency link: Empirical evidence. *Language Awareness*, 17(3), 229-248.
- Guion, S. G., Flege, J., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, 107(5), 2711-2724.
- Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the Critical Period hypothesis for second-language acquisition. *Psychological Science*, 14(1), 31-38.
- Loup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Re-examining the Critical Period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition*, 16(1), 73-98.

- Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*, 24(1), 131-161.
- Johnson, J., & Newport, E. (1991). Critical Period effects on universal properties of language: The status of subadjacency in the acquisition of a second language. *Cognition*, 39(3), 215-258.
- Johnson, J., & Newport, E. (1989). Critical Period effects in second language learning: The influence of maturational state on the acquisition of ESL. *Cognitive Psychology*, 21(1), 60-99.
- Kessinger, R. H., & Blumstein, S. E. (1998). Effects of speaking rate on Voice-Onset Time and vowel production: Some implications for perception studies. *Journal of Phonetics*, 26(2), 117-128.
- Klein, M. (2016). Aspir(at)ing to speak like a native: Tracking Voice Onset Time in the acquisition of English stop. *Gmu.edu*, 4(4), 131-136.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford, UK: Pergamon.
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann-Arbour, MI: University of Michigan Press.
- Lin, C., & Wang, H. (2011). Automatic estimation of Voice-Onset Time for word-initial stops by applying random forest to onset detection. *The Journal of the Acoustical Society of America*, 130(1), 514-525.
- Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustic measurements. *International Linguistic Association*, 20(3), 384-422.
- Lisker, L., & Abramson, A. (1967). Some effects of context on Voice Onset Time in English stops. *Language and Speech*, 10(1), 1-28.
- Macken, M., & Barton, D. (1978). The acquisition of the voicing contrast in English: A study of Voice Onset Time in word-initial stop consonants. *Journal of Child Language*, 7(1), 41-74.
- Major, R. C. (1987). English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics*, 15, 197-202.
- McAllister, R., Flege, J., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, 30(2), 229-258.
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25(3), 481-520.
- Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, 21(1), 81-108.
- Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44(1), 1-35.
- Munro, M. J. (1993). Production of English vowels by native speakers of Arabic: Acoustic measurements and accentedness rating. *Language and Speech*, 36(1), 39-66.
- Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49(1), 285-310.
- Murray, T. E. (1995). *The structure of English: Phonetics, phonology, morphology*. Boston, MA: Allyn and Bacon.
- Newman, D. (2002). The phonetic status of Arabic within the world's languages: The uniqueness of the Lughat Al-Daad. *Antwerp Papers in Linguistics*, 65-75.
- Patkowski, M. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics*, 11(1), 73-89.
- Piske, T., MacKay, I. R. A., & Flege, J. (2001). Factors affecting degree of foreign accents in an L2: A review. *Journal of Phonetics*, 29, 191-215.
- Riney, T. J., & Takagi, N. (1999). Global foreign accent and Voice Onset Time among Japanese EFL speakers. *Language Learning*, 42(2), 275-302.
- Sabir, I., & Alsaeed, N. (2014). A brief description of consonants in modern Standard Arabic. *Linguistics and Literature Studies*, 2(7), 185-189.
- Saito, K. (2015a). Experience effects on the development of late second language learners' oral proficiency. *Language Learning*, 65(3), 563-595.
- Saito, K. (2015b). The role of age of acquisition in late second

- language oral proficiency attainment. *Studies in Second Language Acquisition*, 37(4), 713-743.
- Saito, K., & Brajot, F-X. (2013). Scrutinizing the role of length of residence and age of acquisition in the interlanguage pronunciation development of English /r/ by late Japanese bilinguals. *Language & Cognition*, 16(4), 847-863.
- Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the Critical Period for human speech*. New York, NY: Newbury House/Harper & Row.
- Scovel, T. (1995). Differentiation, recognition, and identification in the discrimination of foreign accents. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 169-182). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Singleton, D., & Ryan, L. (2004). *Language acquisition: Age factor*. Clevedon, Buffalo, Toronto: Multilingual Matters LTD.
- Stevens, G. (1999). Age at immigration and second language proficiency among foreign-born adults. *Language in Society*, 28(4), 555-578.
- Stölten, K., Abrahamsson, N., & Hyltenstam, K. (2014). Effects of age of learning on Voice Onset Time: Categorical perception of Swedish stops by near-native L2 speakers. *Studies in Second Language Learning*, 37(1), 471-100.
- Yeni-Komshian, G. H., Caramazza, F., & Preston, M. S. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics*, 5, 35-48.
- Yu, V. Y., De Nil, L. F., & Pang, E. W. (2015). Effect of age, sex and syllable number on Voice Onset Time: Evidence from children's voiceless aspirated stops. *Language and Speech*, 58(2), 152-167.
- Zobl, H. (1980). The formal and developmental selectivity of L1 influence on L2 acquisition. *Language Learning*, 30(1), 43-57.