International Journal of English Linguistics; Vol. 13, No. 2; 2023
ISSN 1923-869X E-ISSN 1923-8703
Published by Canadian Center of Science and Education

Acquisition of Voice Onset Time for Voiced Plosives of English by Adult Learners of Balochistan

Nasir Abbas Syed[1], Sultan Melfi Aldaihani[2] & Shah Bibi[1]

[1] Department of English Language and Linguistics, Lasbela University of Agriculture, Water and Marine Sciences, Balochistan

[2] The English Department, College of Basic Education, Ardhiya, Kuwait

Correspondence: Sultan Melfi Aldaihani, The English Department, College of Basic Education, Ardhiya, Kuwait.

Received: November 30, 2022 Accepted: January 16, 2023 Online Published: February 7, 2023

doi:10.5539/ijel.v13n2p21 URL: https://doi.org/10.5539/ijel.v13n2p21

Abstract

This study focuses on two experiments conducted with Eastern and Western Balochi speakers. In Eastern Balochi, voiceless stops have aspirated features, but in Western Balochi, they are unaspirated. Eighty-four native speakers of both dialects of Balochi participated in this study. Participants of the first experiment produced words of their L1 in a picture naming task, and those of the second experiment read words in English. VOTs of the L1 and L2 voiced stops elicited from recordings of the productions. Results show that speakers of Western Balochi transfer their L1 negative VOTs to L2 English-voiced stops. However, Eastern Balochi speakers produce English-voiced stops with VOTs, significantly different from their L1 VOTs. Though they could not produce English-voiced stops with native-like accuracy, they produced English stops with significantly longer pre-voicing duration than their L1-voiced stops. Therefore, the study concludes that speakers of those languages with stops with negative VOT ranges face more difficulty acquiring L2 voiced stops of short-lag positive VOTs than those learners whose L1 does not have such stops. The speech learning model is used in this study to analyze all results.

Keywords: voiced, voice onset time, adult learner, English

1. Introduction

1.1 Introduction and Background

Languages of the world have different laryngeal settings for the stops. These settings are defined through phonological features like voicing or aspiration. Both are acoustically defined using voice onset time (VOT), along with some other paradigms, as standard correlates for the classification of plosives (Abramson & Whalen, 2017). There are different ranges of VOT for stops in the world languages. Long lag VOT is an acoustic correlate of aspirated stops, and post-burst short lag VOT defines voiceless unaspirated stops (Cho & Ladefoged, 1999). Voiced stops have different VOT ranges depending on the laryngeal setting of languages. Some languages have voiced stops produced with negative VOT, while others have voiced stops with short-lag positive VOT. Languages like German and English have positive VOT for voiced stops, but those like French, Arabic, Japanese, etc., have voiced stops with negative VOT. VOT is measured in negative values when the vocal fold vibration starts before the burst of stops, and positive VOT means the voicing of the vowel following a stop starts shortly after the burst. The languages which have pre-burst lead voicing are called (truly) voicing languages, and those which have short-lag post-burst voiced stops are called aspiration languages (Honeybone, 2005; Iverson & Salmons, 1995).

Pakistani English (PakE) has developed under the influence of the indigenous languages of Pakistan during the last seven decades (Rahman, 2020). Most of Pakistan's languages are voicing since they have truly voiced stops produced with lead-voicing. Resultantly, under the influence of L1s, English-voiced stops are also produced generally with negative VOT by Pakistani learners (Syed, Ansari, & Gopang, 2017). Previous research has studied the acquisition of English-voiced stops by speakers of voicing languages. Most of the previous studies have concluded that the acquisition of voiced stops of aspiration languages is problematic for adult learners who speak truly voicing languages. The current study is a step forward in this direction.

In this connection, it is relevant to point out that Arabic stops are also pre-voiced (Aldaihani, Alhouti, & Alenezi, 2022). Kuwaiti Arabic and Saudi Arabic also have negative VOT. Therefore, we expect similar behavior from adult learners and speakers of English in Pakistan and the Arab world. Many Balochi speakers are living in the Arabian gulf countries, including Kuwait.

The focus of this study is only on the acquisition of voiced stops of English by speakers of two dialects of Balochi. Balochi is a dominant language of the Iranian branch of the Indo-European family. There are three major Balochi dialects: Southern Balochi, Eastern Balochi, and Western Balochi. Each has further sub-dialects. There are some structural variations among these dialects (Jahani, 2019). These variations are different in grammatical features. This study is based only on plosive consonants. Two dialects of Balochi are selected for this study because they vary in the realization of plosives in terms of VOT. Eastern Balochi speakers seem to have aspirated plosives, whereas Southern and Western Balochi speakers have unaspirated stops (ibid). Since both Southern Balochi and Western Balochi speakers are basically in the west of the Eastern Balochi region, in the current paper, we use the term Western Balochi as a cover term for western and southern Balochi. They both have unaspirated stops.

In the current research, we shall study the acquisition of English-voiced stops by adult Balochi speaker learners of English. The results will be analyzed in light of predictions of the speech learning model (Flege, 1995). The speech learning model (SLM) is an instant model evolving based on various researchers' empirical studies. Recently, Flege and his colleagues have revised it in light of the findings of their latest experimentation (Flege, Aoyama, & Bohn, 2021). The latest version is called SLM-r or 'speech learning model revised'. However, we shall use the standard term SLM for the model without discriminating between its different versions. One of the main postulates of SLM, which is of vital relevance to this study, is the view that L1 and L2 phonemic inventories consist in the mind of adult learners in the same mental space. Therefore, adult learners keep adjusting and improving phonetic representations for the categories of sounds in their minds throughout life. We shall see how this postulate of SLM helps us in accounting for learning issues of adult L2 learners of English.

1.2 Significance of the Study

Research is scarce, particularly on regional languages in applied linguistics. There are few studies of Balochi languages based on laboratory analysis using the latest computational techniques. The current study aims to analyze the sounds of a primary Balochistani language, i.e., Balochi, and compare it with the corresponding English sounds using the latest available software (PRAAT). Therefore, through this analysis, the study plans to contribute simultaneously to the field of theoretical linguistics and ELT. The findings of this study will expose the problems adult learners of English face in acquiring the correct pronunciation of stops of English. Most adult L2 learners of English face problems acquiring aspiration and voicing in stops (Simon, 2009). Therefore, the current study suggests steps to help learn aspiration contrast in English. A large population of English learners may benefit from this research's findings.

Although it is noted in the previous literature that voiceless stops are produced with aspiration in Eastern Balochi and without aspiration in Western Balochi, no previous research has recorded VOT ranges of various dialects of Balochi. The current project fills this gap in the literature and records the VOT ranges of two dialects of Balochi. It also studies the influence of L1 VOT ranges on the acquisition of L2 VOT, focusing on the acquisition of voiced stops in English.

2. Literature Review

2.1 Previous Studies Relevant to the Current Topic

Most early SLA studies claim that it is challenging for speakers of voicing languages to acquire voiced stops of aspirating languages. The reason identified for this difficulty is that pre-voicing languages have voiced stops with lead voicing, but the aspirating languages have the same with positive VOT. Thus, a categorical shift is required in the acquisition of such stops. Typically, a categorical shift in L2 learning is considered easier than a gradient change, as per the researchers' claim in second language acquisition (Flege et al., 2021). However, suppose the L1 of a learner has a phoneme in the short-lag positive VOT range. In that case, there is a strong probability of overlap and confusion between L2 voiced stops (with positive VOT) and L1 voiceless unaspirated stops. To avoid this, speakers of voicing languages produce voiced stops of aspirating language negative VOT like the L1 voiced stops. That has already been known about Arabic, Japanese, and Dutch speakers (Simon, 2009). For example, English-voiced stops are produced with positive VOT as voiceless unaspirated stops by native speakers. Therefore, speakers of voicing languages face extreme difficulty in learning English-voiced stops. Studies by Nasukawa (2010), Shimizu (2011), and Simon (2011), among others, have already demonstrated such literature. Native speakers of English differentiate between voiced and voiceless unaspirated

stops of English based on other acoustic correlates or the complementary distribution between these two phonemic categories in many contexts. Such complementary distribution only exists in some languages like Urdu, Balochi, and Arabic. Therefore, L2 learners of Pakistan, Japan, and the Arabian Gulf, including Kuwait, face difficulty developing two different phonemes in the same acoustic space as native speakers of English do. As a result, they equate L2 voiced stops with positive VOT with their L1 pre-voiced stops.

However, in the current decade, the latest research demonstrates that though the acquisition of voiced stops of aspirating languages is complex, it is possible for speakers of voicing languages. In this regard, many studies claim that such learning is possible. For example, Hamzah, Madbouly, Halim, and Abdullah (2020) studied the acquisition of English labial /p-b/ and velar stops /k-g/ to determine if Arabic learners of English can maintain voicing contrast in different places of articulation. The L1 (Arabic) does not have these contrasts, and voiceless velar and coronal stops in Arabic are produced with aspiration in most Arabic dialects (Kulikov, 2020). Hamzah et al. (2020) found that most of their subjects had acquired accurate VOT for English voiced and voiceless stops, although some produced these stops with abnormal VOT ranges. These results support the claim by SLM that a new contrast is easier to acquire for adult L2 learners.

Another critical point raised by the researchers is that cross-linguistic influence is vital in such learning. Schwartz (2020) acoustically analyzed the stops of Polish learners of English to determine the cross-linguistic influence on the learners' productions. The participants undergoing intensive training sessions related to English language learning reflected more substantial cross-linguistic influence in their VOT ranges of voiced than voiceless stops.

Younes and Gathercoleb (2020) administered a picture naming task with fourteen Spanish-English bilinguals and twelve English monolinguals to study VOT of voiced stops of English /b d g/. The results show that labial /b/ had significantly longer lead-voicing than /d/ or /g/ in the speech of Spanish-English bilinguals. Secondly, they exhibited cross-linguistic influences in the production of voiced stops. The influence of language use was also evident as their productions were more like the corresponding phonemes of the language (English), which was dominant in their daily conversation.

Liu and Lin (2021) investigated how Japanese and Russian stops are produced and perceived by adult Chinese learners. For this purpose, they studied the performance of 39 Chinese whose mother tongue was Mandarin, English was their second language, and they were either learning Japanese or Russian as a third language. The researchers used identification and reading task for L3 and only reading tasks for L2. The findings demonstrate that the participants could perceive a new phonetic feature, i.e., pre-voicing in the target languages but could not accurately produce it.

Aldaihani, Al-Houti and Alenezi (2022) studied the acquisition pattern of Learning Proper VOT of English Velar Stops by Kuwaiti Undergraduate Students from Optimality Theory Perspectives. They focused on adult learners' perception and production of English voiced and voiceless velar plosives /g/ and /k/. Adult Kuwaiti learners of English have developed better comprehension of voiceless velar, and their production of voiceless stops was also good, though not native-like. The researchers used identification and reading task for L3 but only reading tasks for L2. The varying performance of the participants was noticed in this regard. Some were good, some were weak, and some were very poor in perception and production of /g/. In this regard, their significant difficulty was controlling negative transfer from their L1. Kuwaiti Arabic has voiced stops produced with negative VOT, whereas English /g/ is produced with positive post-burst VOT. Most of the participants produced English /g/ in a native-like manner. However, some of them produced it with negative VOT, and they also perceived it as /k/ because the token used for perception was produced with short-lag VOT by a native speaker of English. Overall, the performance of the Kuwaiti learners of English was better on /k/ than on /g/.

The above studies highlight that speakers of voicing languages can learn voiced stops of English which are produced with post-burst short-lag VOT. In this regard, the role of cross-linguistic influence is essential. Another essential thing in this connection is that most of the L1s of learners in the above research were languages that have pre-voiced and voiceless unaspirated stops in their phonemic inventory. Such a laryngeal setting makes learning difficult since learners may confuse voiced stops with voiceless unaspirated stops by producing the former with positive short-lag VOT, whereas the same VOT range they also have for voiceless unaspirated stops of their L1. Arabic has a series of voiced aspirated and pre-voiced stops. Thus, the slot for a short-lag VOT range is empty in their inventories which learners can occupy with this linguistic context. Perhaps this is the reason that some Arabic learners of English could successfully acquire voiced stops of English which usually are difficult to acquire for adult learners (Aldaihani, Alhouti, & Alenezi, 2022; Hamzah et al., 2020).

2.2 Research Questions and Hypotheses

The target learners in the current study are native speakers of Balochi, which has voiceless aspirated stops in one dialect (Eastern Balochi) and voiceless unaspirated stops in the other dialect (Western Balochi). Balochi voiced stops are truly voiced stops in both dialects under study. The study hypothesizes that native speakers of Eastern Balochi will be better able to acquire positive VOT values for L2 English-voiced plosives because the short-lag VOT slot is empty in this language. The study aims to answer the following research questions.

1) What are the VOT values of Eastern and Western Balochi?

2) Do L1 VOT ranges have any positive or negative cross-linguistic influence on the acquisition of L2 VOT ranges? If so, what is that effect in the context of Baloch speaker learners of English?

Based on the brief review of previous studies, Western Balochi learners of English may face difficulty in the acquisition of accurate VOT ranges for English-voiced stops. However, speakers of Eastern Balochi may be experiencing relatively less difficulty acquiring these consonants. A detail of the data collection process is described below.

3. Methods of Data Collection

3.1 Stimuli for Data Collection

The data for this study were collected in two phases. There were two groups of participants in each phase. The stimuli for the first phase were four monosyllabic words of Balochi, which start with plosives. All stimuli had low vowel /a/ as nuclei and were commonly used and understood in both dialects of Balochi. The selected stimuli had low vowel /a/ immediately located after stops. It was on purpose because in the previous research, /a/ found to be the least influencing adjacent vowel (Best, McRoberts, & Sithole, 1988; Guion, Flege, Akahane-Yamada, & Pruitt, 2000). The list of stimuli is given below. Each stimulus had three repetitions in the recording of this list.

/ba:g/ ‘garden’, /ɖa:r/ ‘wood’ /ɖ̣:al/ ‘pulses’ /ga:r/ ‘lost’.

The data were collected through a picture naming task in this phase. Participants saw the pictures of the above items and named the items loudly. The productions were recorded. Before recording, it was ascertained that native speakers of both dialects of Balochi were well-familiar with these words of their mother tongue.

The stimuli for the second phase were three English mono-syllabic words starting with voiced plosives. Each target stimulus had three repetitions in the list. Some distracters were included at the beginning, end, and between the target words. The target stimuli were English words ‘bark, dark, guard’. Participants read these words, and their productions were recorded for further analysis.

The reason for using the picture naming task for collecting data from the first group is that the participants of that group were uneducated monolinguals who could not read. Therefore, they were asked to name the pictures. On the other hand, the second group comprised university students who were educated, so a word reading task was developed for them for data collection.

3.2 Details of Participants

There were two groups of participants in the first phase of data collection, one comprising twenty-four male monolingual native speakers of Eastern Balochi (mean age = 27.00 years, st.dev. = 8.40, range = 16–30) living in Balochistan. The other group was of twenty male native speakers of Western Balochi (mean age = 25.75 years, st. dev. = 3.78, range = 20–35) living and working in Kuwait.

In the second recording phase, twenty Western Balochi speaker learners (mean age = 21.20 years, st. dev. = 2.14, range: 18–26) and twenty Eastern Balochi speakers (mean age = 24.05 years, st. dev. = 3.76, range = 18–29) learners of English participated. The second experiment participants were also all male students studying in the third semester of their undergraduate degree courses. All experiment participants were selected based on availability and convenience sampling. All learner participants had studied English for more than a year at the undergraduate level. Before joining the university, they also studied English at schools and colleges. Therefore, these participants had a long experience learning academic English.

3.3 Tools and Methods of Data Collection

M-Audio digital recording device was used for the recording of participants. The experiments with learners were conducted in the computer labs of a university. Western Balochi Monolingual Participants were recorded in Kuwait, and those of Eastern Balochi in Pakistan. The recordings obtained in both phases of data collection were analyzed acoustically using Praat (Boersma & Weenink, 2019) in prosody-pro (Xu, 2021) software. SPSS 18

was used for quantitative data analysis and testing of the hypothesis.

The whole procedure of recording was explained to the participants in both phases of experiments, although the specific objectives of data collection were concealed from them. None of the participants reported any hearing or speaking loss. They participated in this study voluntarily. The data is presented and discussed in the following sections.

4. Presentation of Data

The data are presented in this section. In the following tables, VOTs of voiced stops of native speakers of Eastern Balochi and Western Balochi (L1s) stops and those of English (L2) voiced plosives produced by learners are compared. An independent sample t-test was applied to the mean scores of the L1 and L2 data. T-tests were applied to determine the significance of the difference between groups. The results of the tests are quoted in the last column of each table. The L1 data is based on two coronal plosives of Balochi, namely, retroflex /ɖ/ and dental /d̪/ but corresponding to these, English has a single /d/ at the alveolar place of articulation. Therefore, both L1 coronal stops are compared with the sole voiced alveolar stop of English in each table.

Table 1. Comparison of VOTs of Eastern Balochi and L2-English voiced stops produced by Eastern Balochi learners

L1	Mean (Standard Deviation)	L2	Mean (Standard Deviation)	Significance
/b/	-39.57 (24.00)	/b/	-68.00 (14.46)	t= 4.64, p=.001
/ɖ/	-35.95 (21.99)	/d/	-67.01 (21.76)	t=4.64, p=.001
/d̪/	-31.30 (23.72)	/d/	-67.01 (21.76)	t=5.16, p=.001
/g/	-33.44 (14.88)	/g/	-56.49 (21.05)	t=4.243, p=.001

The above results show significant differences between VOT ranges of voiced stops of the L1 (Eastern Balochi) and L2 (English). It means Eastern Balochi speaker learners of English have developed two separate VOT ranges for L2 and L1 voiced stops. The results show learners have produced VOTs of English stops with relatively longer pre-voicing duration. In other words, they have attempted to produce English-voiced stops with a VOT range that is deflected away from the native English VOT range.

The following table presents VOT values of L1 and L2 stops elicited in the speech of Western Balochi speaker learners. The results of the t-test applied on L1 Western Balochi voiced stops and L2 English voiced stops are also given in the last column.

Table 2. Comparison of VOTs of Western Balochi and L2-English voiced stops produced by Western Balochi learners

L1	Mean (Standard Deviation)	L2	Mean (Standard Deviation)	Significance
/b/	-85.05 (21.00)	/b/	-71.84 (33.36)	t=-1.476, p=0.150
/ɖ/	-78.74 (22.52)	/d/	-71.00 (28.00)	t=-0.984, p=0.331
/d̪/	-73.43 (33.05)	/d/	-71.00 (28.00)	t=-0.264, p=0.793
/g/	-62.38 (23.00)	/g/	-66.46 (28.00)	t=0.515, p=0.610

The difference between the means of voiced plosives of the L1 (Western Balochi) and L2 (English) produced by monolingual native speakers of Western Balochi and Western Balochi learners of English are non-significant. This indicates a negative transfer from the L1 into the L2. We analyze these results in light of the speech learning model (SLM) in the next section.

5. Analysis and Discussion

The results show that voiced stops of both dialects of Balochi are articulated with negative VOT. It is already known that most Pakistani languages have voiced stops with negative VOT (Syed, 2014). Balochi is also similar to other Pakistani languages in this regard. According to the classification of languages by Honeybone (2005), Balochi, like Arabic, is a voicing language, whereas English is a language that discriminates between stops based on aspiration.

Voiced stops of English by adult learners are articulated with negative VOT by speakers of both dialects. Like other languages, Balochi speakers produce labial voiced stops with longer pre-voicing duration than velar stops. Kuwaiti Arabic learners have also been found to have the same (Aldaihani, Alhouti, & Alenezi, 2022). The

results of the t-test show no significant differences between stops of L1 Western Balochi and those of L2-English. This shows that participants have exploited their L1 laryngeal contrast in the articulation of voiced stops of English and have yet to develop a new phonetic category for these consonants of English. This situation can be explained in the Speech Learning Model (Flege, 1995; Flege et al., 2021). According to one of the assumptions of the SLM, if learners perceive an L2 phoneme as the closer L1 phoneme, there is a strong probability of interference of the L1 in the perception and production of the L2 phoneme. This is called 'equivalence classification' between sounds (Flege, 1987). Equivalence classification between sounds develops because, as per SLM postulates, both L1 and L2 sounds exist in the same mental space of adult learners. It is already known in the literature that equivalence classification between L1 and L2 sounds does not allow the formation of a new phonetic category for an L2 sound and leads to negative transfer (Chang, 2019; Leeuw, Tusha, & Schmid, 2018). Thus, we infer from these results that Western Balochi learners of English consider English-voiced stops identical to their L1-voiced stops and develop a kind of equivalence classification between the L2 and the corresponding L1 sounds, which blocks their learning.

The learners who speak Eastern Balochi do not have unaspirated stops in their L1 phonemic inventory. Therefore, they can easily acquire unaspirated stops of English. The current group of learners was also able to produce unaspirated stops of English successfully in another experiment. As voiceless unaspirated stops occupy the short-lag VOT space, they cannot allocate the same space to the voiced stops. Thus, they deflected their L2 voiced stops towards the other end and produced them with a longer pre-voicing duration. It is also a kind of development. Such development is normally observed when learners realize some difference between an L2 sound and the closest L1 sound, but they cannot acquire such an L2 sound accurately (Mooney, 2019; Syed & Saleem, 2019).

What practically seems to happen in this context is that the Eastern Balochi group of L2 learners attempts to develop a sound category that has a VOT range different from the L1 VOT range. Thus, they drift their target L2 VOT range for the voiced stop of English. However, they have already developed a new phoneme category in the target range of voiceless unaspirated stops for English voiceless stops. Thus, their new category for English-voiced stops could not reach the target level of the post-burst positive VOT range but significantly drifted to the other direction far from the L1 VOT range. The results confirm that those learners whose L1 has voiceless aspirated stops but no voiceless unaspirated stops (i.e., with an empty slot in their L1 phonemic inventory for stops with short-lag VOT range), are able to differentiate between pre-voiced and short-lag VOT stops.

The overall results show that in voiced stops, mostly L1 laryngeal contrast is transferred through one group of participants (Eastern Balochi) has yielded slightly different results. The results yielded by the Western Balochi group in this study follow most of the previous studies of the first decade of the current century, which claim that short-lag VOT for voiced stops of aspirating languages (like English) is difficult to learn for speakers of the voicing languages (Flege & Port, 1981; Nasukawa, 2010; Shimizu, 2011; Simon, 2009). However, the results obtained from the Eastern Balochi group of learners support the findings of Hamzah et al. (2020) and Aldaihani, Alhouti and Alenezi (2022), which confirms that adult learners whose L1 (like Arabic) has voiced stops with negative VOT. While voiceless stops with long lag VOT (and with an empty slot for short-lag VOT) are better able to perceive problems &in the acquisition of voiced stops of aspirating languages (like English) produced with short-lag post-burst VOT. This situation warrants another study to confirm the assumption that Eastern Balochi learners of English have a better perception of voiced stops of English because, in the current study, we only recorded the production of participants. However, we know that perception and production correspond to L2 acquisition (Felge, 1995).

6. Conclusion

This study was conducted with speakers of Balochi who were learning English at a university. The main aim of the study was to measure VOT ranges of stops in two dialects of Balochi and their cross-linguistic effect on L2 VOT ranges of English voiced stops. Experiments were conducted with the participants in two different places. In each phase, participants comprised native speakers of Eastern Balochi and Western Balochi. The first experiment was based on a picture naming task. The participants saw the pictures in this task and produced words starting with voiced stops. These productions were recorded and analyzed acoustically. The acoustic analyses show that Balochi voiced stops are produced with negative VOT ranges. The second experiment was also conducted with two groups of adult learners of English who were also native speakers of Eastern Balochi and Western Balochi. The second experiment was based on a word-reading task. In both experiments, the vocalic context was controlled by selecting stimuli of similar nuclei. The results of the second experiment show that Western Balochi speakers transferred L1 VOT ranges to L2-English voiced stops; no significant difference was

found between their L1 and L2 pre-voicing. However, results obtained from Eastern Balochi speakers show some development in that there was a significant difference in means between the L1 and L2 voiced stops. Their L2 VOT range, though still negative, significantly drifted away toward the positive VOT range. Thus, the results confirm that learners whose L1 has aspirated and pre-voiced stops only, with an empty slot for stops with short-lag VOT, can realize the difference between the L1 pre-voiced stops and L2 voiced stops of short-lag positive VOTs with relative ease as compared to those learners whose L1 has short-lag VOT stops. In this way, the study also confirms the interaction between L1 and L2.

Author contribution

The first author conceived the main idea of this paper. The second and third authors collected data in Kuwait and Balochistan, respectively. The second and third authors also did acoustic analyses. Overall, theoretical analysis and write-up were completed by all authors together.

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

Aldaihani, S. M., Al-Houti, S. K., & Alenezi, Y. M. (2022). Learning proper VOT of English velar stops by Kuwaiti undergraduate students from VOT and Optimality Theory perspectives. *International Journal of Education, Culture and Society*, *7*(1). https://doi.org/10.11648/j.ijecs.20220701.16

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345–360. https://doi.org/10.1037/0096-1523.14.3.345

Boersma, P., & Weenink, D. (2019). *Praat: doing phonetics by computer*.

Chang, C. B. (2019). The phonetics of second language learning and bilingualism. In *The Routledge Handbook of Phonetics* (pp. 427–477). Abingdon: Routledge. https://doi.org/10.4324/9780429056253-16

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207–229. https://doi.org/10.1006/jpho.1999.0094

Flege, J. E. (1987). The production of 'new' and 'similar' phones in foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*, 47–65. https://doi.org/10.1016/S0095-4470(19)30537-6

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). New York: Timonium, MD.

Flege, J. E., Aoyama, K., & Bohn, O. S. (2021). The Revised Speech Learning Model (SLM-r) Applied. In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3–83). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108886901.002

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language & Speech*, *24*(2), 125–146. https://doi.org/10.1177/002383098102400202

Guion, S. G., Flege, J. E., 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. https://doi.org/10.1121/1.428657

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

Honeybone, P. (2005). *Diachronic evidence in segmental phonology: The case of laryngeal specifications*. Berlin: Mouton de Gruyter.

Iverson, G. K., & Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, *12*(3), 369–396. https://doi.org/10.1017/S0952675700002566

Jahani, C. (2019). *A Grammar of Modern Standard Balochi*. Uppsala: Uppsala University.

Kulikov, V. (2020). Laryngeal Contrast in Qatari Arabic: Effect of Speaking Rate on Voice Onset Time.

Phonetica, *77*(3), 163–185. https://doi.org/10.1159/000497277

Leeuw, E. D., Tusha, A., & Schmid, M. (2018). Individual phonological attrition in Albanian-English late bilinguals. *Bilingualism: Language and Cognition*, *21*(2), 278–295. https://doi.org/10.1017/S1366728917000025

Liu, J., & Lin, J. (2021). *A Cross-Linguistic Study of L3 Phonological Acquisition of Stop Contrasts*. SAGE Open. https://doi.org/10.1177/2158244020985510

Mooney, D. (2019). Phonetic transfer in language contact: Evidence for equivalence classification in the mid-vowels of Occitan-French bilinguals. *Journal of the International Phonetic Association*, *49*(1), 53–85. https://doi.org/10.1017/S0025100317000366

Nasukawa, K. (2010). Place-Dependent VOT in L2 Acquisition. In M. T. P. e. al (Ed.), *Selected Proceedings of the 2008 Second Language Research Forum* (pp. 198–210). Somerville: Cascadilla Proceedings Project.

Rahman, T. (2020). Pakistani English. In K. Bolton, W. Botha & A. Kirkpatrick (Eds.), *The Handbook of Asian Englishes* (pp. 279–296). Hoboken, NJ: Wiley Blackwell. https://doi.org/10.1002/9781118791882.ch11

Schwartz, G. (2020). Asymmetrical cross-language phonetic interaction—phonological implications. *Linguistic Approaches to Bilingualism*, 1–30. https://doi.org/10.1075/lab.19092.sch

Shimizu, K. (2011). *Study on VOT of initial stops in English produced by Korean, Thai and Chinese speakers as L2 learners*. Paper presented at the. ICPhS XVII, Hong Kong.

Simon, E. (2009). Acquiring a new second language contrast: an analysis of the English laryngeal system of first language Dutch speakers. *Second Language Research*, *25*(3), 377–408. https://doi.org/10.1177/0267658309104580

Simon, E. (2011). Laryngeal stop systems in contact: Connecting present-day acquisition findings and historical contact hypotheses. *Diachronica*, *28*(2), 225–254. https://doi.org/10.1075/dia.28.2.03sim

Syed, N. A. (2014). Why Pakistani learners cannot acquire voiced stops of English? *Journal of Research and Social Sciences*, *1*(1), 19–30.

Syed, N. A., Ansari, S., & Gopang, I. B. (2017). Perception and Production of Consonants of English by Pakistani Speakers. *International Journal of English Linguistics*, *7*(3), 201–214. https://doi.org/10.5539/ijel.v7n3p201

Syed, N. A., & Saleem, A. (2019). A Review of the Speech Learning Model in the Perspectives of Learners of English in Pakistan. *Journal of Independent Studies and Research*, *17*(2), 85–110. https://doi.org/10.31384/jisrmsse/2019.17.2.12

Xu, Y. (2021). *ProsodyPro—A Tool for Large-scale Systematic Prosody Analysis of Continuous Prosodic Events*.

Younes, S. A., & Gathercoleb, V. C. M. (2022). Phonological interaction in Spanish-English bilinguals: effects of cognate usage on voice onset time of voiced stops. *International Journal of Bilingual Education and Bilingualism*, *25*(2), 709–727. https://doi.org/10.1080/13670050.2020.1715914