10.31384/jisrmsse/2019.17.2.12

A Review of the 'Speech Learning Model' in the Perspectives of Learners of English in Pakistan

Nasir Abbas Syed[1*]
Arshad Saleem[2]

ABSTRACT

This paper reports an experiment in which production of English plosives by advanced Pakistani learners was studied in the light of predictions of speech learning model (SLM). The results show that the learners equate aspirated and unaspirated allophones of English labial and coronal stops with the unaspirated stops of their L1. However, they maintain a difference between two allophones of English velar stop. Some participants also increase VOT of the unaspirated allophones of English stops to a level which is larger than the native VOT range. The directionality of learning is from dorsal to labial to coronal. The acquisition of accurate VOT of English coronal stops is found to be more difficult for Pakistani learners because they substitute English alveolar stops with retroflex stops of their L1. The findings of this study also points out some gaps in the SLM on the basis of which revisions are suggested in the SLM.

JEL Classification: *D8, D80, D83*

Keywords: *Acquisition, consonant, L2, perception, SLM, VOT*

INTRODUCTION

Acquisition of phonemes of a second language has been a point of much discussion and debate among researchers for a number of decades. The contrastive analysis hypothesis (hereafter CAH) by Lado (1957) was a turning point in this debate (Wardhaugh, 1970). The CAH claims that differences between L1 and L2 cause errors in second language acquisition. According to the CAH, phonemes of an L2, which are different from those of L1, may be difficult to acquire. Thus, fundamental concept of the CAH was based on a contrastive analysis of the phonemic inventories of L1 and L2. Many other models of second language acquisition have been presented after the CAH. Some of these challenged predictions of the CAH with respect to difficulty/ease of learning, but almost all of them focus on a contrastive analysis of the phonemic inventories of L2 and L1. A prominent factor discussed by almost all popular models of second language acquisition is interference of L1 in L2 acquisition. These models, by some means or others, seem to agree that errors in acquisition of L2 sounds are because of L1 of learners.

Structural conformity hypothesis (SCH) (Eckman, 1991) is the only model which predicts directionality of learning on the basis of markedness of L2 structure. According to the SCH, acquisition of a CCC structure is more difficult than that of a CC structure, a subset of the CCC structure, and the CC structure is more difficult to acquire than a single consonant, a subset of the CC structure. Thus, the SCH predicts a direction of acquisition from C to CC to CCC. As pointed out in the SCH, another important factor in L2 acquisition is markedness (Eckman, 1977, 1991). The generalization is that more marked structures are more difficult to acquire. Later models such as perceptual assimilation model (hereinafter PAM), feature model (hereinafter FM) and speech learning model (hereinafter SLM) attempt to predict a

1- Lasbela University of Agriculture, Water & Marine Sciences, Lasbela, Pakistan.
2- Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Karachi, Pakistan.
*- Corresponding Author Email: nasirabbasla@gmail.com

directionality of ease/difficulty in learning, with detailed classification of L2 sounds and with more focus on perception rather than production. The SLM (Flege, 1995), PAM (Best, 1994, 1995), and FM (Brown, 1998, 2000) have strongly influenced the research in the field of second language acquisition. The current study was conducted for testing the predictions of the SLM.

The SLM is based on empirical studies conducted by Flege and colleagues (Flege, 1987, 1988, 1992a, 1993, 2009; Flege & Eefting, 1988; Flege & Fletcher, 1992; Flege, Takagi, & Mann, 1996; Flege, Yeni-Komshian, & Liu, 1999). The model has also been tested and found quite accurate in predicting about acquisition of tones of L2 by experienced adult learners (Hao, 2014). Although there is a huge body of literature on the SLM, there is no study on Pakistani learners of English with reference to this model. As Larson-Hall (2004) recommends, for developing wider generalization, these models of second language acquisition be extended to and tested on different groups of learners with varying linguistic and cultural backgrounds. The current study serves the same purpose. On account of its empirical evidence, the study recommends some additions and reviews in the speech learning model.

The remainder of this paper is divided into different sections. The following section briefly introduces and analyses the speech learning model. In section 2, laryngeal settings of the L1 (Saraiki) of the participants of this study and L2 (English) are compared. Section 3 is based on research methodology, which provides detail of participants and the experiment conducted with them. The data are presented in section 4 and discussed and analyzed in section 5. In sections 6, the speech learning model is revised in light of findings of this study. The paper ends with conclusion in section 7.

Speech learning model (SLM)

According to Flege (2003, p. 326), speech learning model (SLM) 'is the only extant theory that focuses explicitly on L2 speech acquisition'. The SLM classifies L2 sounds into three categories, namely 'identical', 'similar' and 'new', of which 'similar' sounds pose the greatest difficulty for adult L2 learners (Fox & McGory, 2007, p. 108). The SLM predicts that learning new L2 sounds is easier than learning sounds which are similar to L1 sounds. The following are main postulates of the SLM reproduced from Flege (1995, p. 239).

1. "The mechanical processes used in learning the L1 sound system including category formation, remain intact over the life span and can be applied to L2 learning.
2. Language-specific phonetic aspects of speech sounds are specified in long-term memory representations called phonetic categories.
3. Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as the realization of each category.
4. Bilinguals strive to maintain contrast between L1 and L2 phonetic categories which exist in a common phonological space."

The SLM is based on an assumption that ability to acquire new sounds remains effective and evolves throughout life (Flege, 1995, p. 239). This is in contrast to the critical period hypothesis which claims that there is a period for acquisition of new sounds after which the ability to acquire new sounds ends or drastically diminishes (Lenneberg, 1967; Patkowski, 1990; Penfield & Roberts, 1959; Scovel, 1988). There is a lot of empirical evidence in the literature to substantiate the CPH (Bley-Vroman, 1990; Clahsen & Muysken, 1986, 1989). But as Larson-Hall (2004) comments, Flege does not believe in a critical period or absolute loss of perceptual abilities that accounts for adult difficulties.

The SLM does not disregard the idea of filtering of L2 material through the L1 at an initial stage of learning; it rather claims that long exposure to an L2 may block such a filtering (Flege, 2003). Flege claims that a phonetic shift definitely occurs between the ages of five and seven which 'may render late learners less able to establish additional phonetic categories for sounds after rather than before the age of five to seven years' (Flege, 1992). Thus, the difference between the CPH and SLM is that, while the former predicts that ability to acquire an L2 sound diminishes around the age of puberty, the latter predicts that the same ability minimizes at the age of five to seven. Another difference between the two theories is regarding the reason for the loss of ability to learn a new language. The CPH attributes the loss of learning new sounds after the critical period, to neurological maturation (Scovel, 1988) which occurs on account of biological development, whereas the SLM assumes that ability to learn L2 sounds diminishes because the L1 phonetic system has already stabilized when adult L2 learners start learning a new language (Flege, 1992).

Flege (1995) argues that if learning-ability diminishes with the end of critical period (age) of learning, there should be a sudden drop in improvement of L2 learners at the end of that period. But many studies have found improvement in adult learners who had started learning an L2 after the age of puberty (see studies quoted in Flege, 1995 for example). Similarly, Flege argues (ibid) that if there is a critical period after which no learning occurs, then all new L2 sounds should be equally difficult for adult learners, but it has been observed that all L2 sounds are not equally difficult for adult learners. Thus, the SLM does not thoroughly accord with the predictions of the CPH.

The SLM mainly focuses on very advanced learners as well as bilinguals. It predicts that phonetic categories of sounds developed by bilinguals may be different from those of L1 speakers (Flege, 1995). This difference, also maintained by advanced L2 learners, is attributed to feature-based and acoustic differences between L1 and L2 sounds leading bilinguals/adult-learners to develop slightly different categories from those developed by monolingual speakers of the target language. An example of such a case (Flege, 1987b) is of French learners of English who developed a VOT range for French (L1) stops which was different from the VOT range of French monolinguals. The VOT of French stops of the bilinguals was between French and English VOT ranges.

Guion et al. (2000) recommend to take extreme care in application of the SLM to the beginners. The basic reason is that this model is developed for advanced learners (Best & Tyler 2007). The model gives following seven hypotheses involving different situations of second language acquisition at advance level of L2 acquisition (Flege, 1995).

H1: The phonemes of L1 and L2 are perceptually related to each other at a position-sensitive phonetic allophonic rather than an abstract phonemic level.

H2: A new phonetic category for L2 sounds may be developed if learners discern a difference between a particular sound of L2 and the closest sound(s).

H3: The discrimination between L2 and L1 phonemes is a function of the perceptual distance between the two sounds.

H4: The age of learning negatively correlates to learning a new language in that, as the former increases the latter decreases.
H5: Equivalence classification between two sounds blocks formation of a separate phonetic category for a new L2 sound.

H6: The phonetic categories developed by bilinguals may be different from those of monolinguals, either for maintaining a contrast between L2 and the corresponding L1 sounds, or if the categories of bilinguals are based on features different from those of monolinguals of the L2.

H7: There is a correspondence between perception and production of sounds by L2 learners.

Although SLM is a model which studies acquisition of L2 sounds in detail, there are still some issues which either have not been addressed in this model, or they need further clarification. The main issue in this regard is absence of a proper method for measuring phonetic similarity or distance between sounds (Harnsberger, 2001, Schmidt, 1996). Flege also comes across these difficulties in testing these hypotheses which he himself describes in the following words (1995):

> "An obstacle to testing hypotheses such as these is the lack of an objective means for gauging degree of perceived cross-language phonetic distance. It is uncertain, also, as what metric bilinguals use in doing so."

In another context, Flege (2003) comments about this issue pointing out that,

> "A limitation of the SLM is that it does not provide a metric for determining when cross-language phonetic differences will be too small to support category formation......"

Although Flege has suggested some ways to address this difficulty, he does not commit to a particular method (Guion et al. 2000). After suggesting some ways to gauge the similarity/distance between two sounds, Flege (1995) leaves the question open. He (1995) thinks that phonological features play a role in the perception of L2 sounds. Another option that Flege (1997) suggests is using phonetic details as represented by IPA symbols. Flege (1992) treats this idea as a 'provisional' solution and suggests 'to supplement the phonetic symbol test with additional acoustic criteria' (Flege, 1997). In the literature, the following measures to gauge the distance between sounds of L1 and L2 have been suggested:

a. Articulatory gestures (Browman & Goldstein, 1990)
b. Perception test (Guion et al., 2000)
c. IPA symbols (Flege, 1997)
d. Phonetic details (Ladefoged, 1990)
e. Phonological features (Flege, 1995)

Despite all this, as pointed out by Kato (2005), SLM still lacks a precise algorithm to determine perceptual distance between L1 and L2 sounds, although its predictions are based on such distance. In the absence of such a precise algorithm, SLM cannot determine level of difficulty between two L2 sounds which are in the same category according to the classification of the model. For instance, the SLM predicts that a similar sound is more difficult than a new one but it cannot determine a level and directionality of difficulty between two new or two similar sounds of an L2.

Related to this is another issue of the classification of L2 sounds as 'new' and 'similar'. According to Flege (1988) "New L2 phones have no counterpart in the L1 and so by definition differ acoustically from phones found in L1" whereas the "similar L2 phones on the other hand, differ systematically from an easily identifiable counterpart in L1." These definitions are vague becuase we do not know when a sound is easily identifiable as similar to the corresponding L1 sound. For the current analysis, we shall consider an L2 sound 'similar' to the closest L1 sound if there is only gradient phonetic difference between the two sounds but both have the same place and manner of articulation features. However, if an L2 sound and the closest L1 sound have different phonological features, both will be considered different and the L2 sound will be treated as 'new'. For example, English (L2 in the current study) /p k/ have the corresponding sounds in Saraiki (L1 in the current study) which have the same place and manner of articulation with only minor phonetic differences of voice onset time (VOT). Therefore, these sounds may be treated as similar for the participants of this study. But English /t/ is different from the closest Saraiki sounds in terms of phonological features which are discussed in detail in the follwoing section. Thus, English /t/ may be treated as a new sound for the participants of this study.

An important criticism of the SLM comes from Larson-Hall (2004), whose empirical study concluded that the SLM is not so helpful in predicting difficulties of L2 learners in acquisition of new sounds of a second language. She concludes, on account of her empirical study, that the FM, rather than the SLM, has stronger predictability about expected outcomes in specific L2 learning situations. The study by Lai (2009) also poses another challenge for the SLM. Lai studied perception of Malay and Burmese learners of Mandarin Chinese in the light of predictions of the SLM. An interesting finding of this study is that similarity effect does not seem to be the only factor which causes difficulty in perception of L2 sounds. The role of markedness was quite apparent on performance of participants. Thus, the author concludes that the L1-L2 segmental inventories (in terms of Flege's classification of L2 sounds into 'similar' and 'new') are not sufficient for understanding the difficulties of L2 learners. Lai (2009) suggests an interactive model of second language acquisition based on articulatory, auditory and markedness factors to explain difficulties of L2 learners.

Lopez (2012) conducted an experiment with 16 adult English learners who started learning Spanish in classrooms at an average age of 13.1 years. They were intermediate level learners at the time of experiment. VOT of English and Spanish stops [p t k] produced in English and Spanish monolingual sentences and in English-Spanish and Spanish-English code switched sentences were taken for analysis. English is an aspiration language (Honeybone, 2005) which has stressed stops with long-lag VOT but Spanish stops are produced with short-lag VOT (Flege & Eefting, 1988). The target stimuli of English and Spanish had stops on word-initial

3- This is called 'merger hypothesis' in the SLM literature (Flege, 1987). An example of such a merger is in Dutch L2 phonemic inventory. They merge English /t/ and /θ/ into a single phonetic category (Wieling, Veenstra, Adank, & Tiede, 2017).

positions. The participants produced English /p/, /t/ and /k/ with 52, 57 and 64 ms respectively and Spanish stops with 27, 27 and 44 ms respectively . The VOT of Spanish stops as noted by Lisker and Abramson (1964) are 4, 9 and 29 respectively. The study was conducted to test the claim of SLM that the mechanism which develops separate phonetic categories remains operative throughout life-span. The results show that the hypothesis was partially tested. Although the adult learners could not develop quite native-like categories of the target sounds, they have developed two separate VOT ranges for these sounds in their L2 phonemic inventories. The interaction between L1 and L2 phonetic categories is also apparent from the results. An important finding of this study is that place of articulation had strong effect on acquisition of the L2 sounds. The results show that adult learners could not suppress aspiration in production of /k/ in Spanish. They produced Spanish /p/ and /t/ with a mean VOT of 27 ms which is within the range of voiceless unaspirated stops (though deviated from the VOT ranges of monolingual Spanish speakers) but they produced /k/ with 44 ms VOT which is a bigger VOT for an unaspirated stop. The findings of this study confirm that L2 learners produce some sounds of L1 and L2 which have overlapping categories with a comprised values of both languages. This further confirms the SLM hyopthesis that L1 and L2 phonemic inventories exist in the same acoustic space. With reference to the results obtained in production of /k/, Lopez comments that it is quite possible that "learners reserve the overlapping area for the production of L2 Spanish /k/ until the L2 phonetic category develops more fully". Overall the findings of the study provide empirical support to the SLM.

Schmid, Gilbers, and Nota (2014) studied ultimate attainment in late second language acquisition of 20 Dutch learners of English with high proficiency in the target language. 9 native English speakers living in the Netherlands were also taken as a control group. The control group participants were English-Dutch bilinguals fully competent in Dutch. The performance of the participants was tested in VOT, vowel formants and ellipsis in the target language (English). The results of VOT analysis only which are of relevance for the currrent study are discussed here. The recording was done in a carrier sentences 'I say ......' and in free a talk in the form of a commentary on a short movie. Accoustic analyses show that VOTs of the L2 learners were similar to the native speakers (control group) in stops produced in carrier sentences as well as in free speech. However, on the basis of overall results the authors conclude that 'even the most dedicated and most successful late L2 learners may encounter some 'pockets' of L2 grammar or phonetics that prove difficult to fully master'. This agian confirms the SLM hyopthesis that categories of L2 sounds developed by adult L2 learners may be a little deflected away from those of monolinguals of the target language[4].

The current study also aims to test the hypotheses of the SLM in a new context which has already not been tested with this view. The study focuses on acquisition of English stops by Saraiki learners of English. VOT will be used as an acoustic correlate to measure the phonetic categories of participants. In the following sub-section, laryngeal settings of Saraiki (L2) and English (L2) are defined briefly.

Laryngeal settings of Saraiki (L1) and English (L2)

The current study is based on an experiment conducted with 32 participants of the control groups (10 Saraiki and 22 English monolinguals) and 60 adult Pakistani learners of English

4- Shmid et al. (2014) are of the opinion that bilingual and monolingual speakers experience two different type of input and mental processes; therefore the achievements of L2 learners should only be compared with those bilingual speakers who are equally competent in the L1 and the target langauge.

who speak Saraiki as L1. Saraiki is an Indo-Aryan language which has aspiration contrast on phonemic level (Shackle, 1976). The difference between English /p t k/ and the corresponding Saraiki sounds is that English has aspiration contrast at allophonic level whereas in Saraiki, the aspirated and unaspirated stops are phonemes. The study is based on voice onset time of English plosives produced by adult learners. The voice onset time was selected as an acoustic correlate to study acquisition of L2 plosives because it provides a solid statistical yardstick to determine distance between two sounds. Voice onset time, a term developed by Lisker and Abramson (1964), is as defined as time interval between the burst of a stop closure and the onset of voicing for a following vowel calculated in milliseconds (Docherty, 1992). In articulation of stops, the active articulator touches the passive articulator, remains in the same position for a while and then suddenly separates for production of the target sound. English has short-lag VOT for voiced and voiceless unaspirated stops, and long-lag VOT for aspirated stops. Saraiki has pre-voicing for voiced stops, short-lag VOT for voiceless unaspirated stops and long-lag VOT for the aspirated voiceless stops. In this way, Saraiki has a four-way laryngeal contrast having voiced/voiceless, aspirated/unaspirated and plain/breathy voiced explosives. Besides, it also has implosive nasal breathy voiced consonants in its phonemic inventory. The language also has plain and breathy voiced nasals (See Shackle, 1976 for detailed description of Saraiki consonant phonemic inventory). Saraiki voiceless stops have almost the same range of VOT as English voiced stops which may create difficulties for Saraiki learners of English in perception of English voiced stops[5]. The following figure adapted from Nasukawa (2010) reflects the VOT ranges of English and Saraiki [t d] stops.

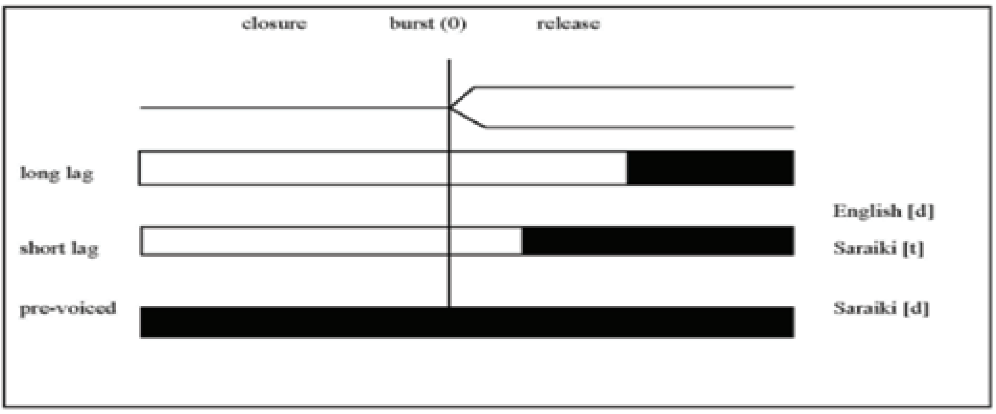

Figure 1: *VOT settings of English and Saraiki*

In terms of VOT Saraiki and English voiceless aspirated stops are the same in both language, but Saraiki unaspirated stops and English voiced stop have almost similar VOT ranges. The current study only focuses on acquisition of voiceless stops of English by adult Saraiki learners

RESEARCH METHODOLOGY

An experiment was conducted with two groups of learners who speak Saraiki as the L1. Two control groups were also among the participants of this study. Ethical approval for the experiment was obtained from University of Essex. The detail of the participants and experiment are given in the following sub-sections.

5- See (Syed, 2013a, 2014) for difficulties of Saraiki learners in acquisition English voiced stops.

Participants

As mentioned above, there are two groups of target learner participants in this study, one from Pakistan and another from the UK. Two control groups also participated in the experiment. The Pakistan-based control group comprised of 10 Saraiki monolinguals and the UK-based control group comprised of 22 English monolingual native speakers. The details of the participants are given in the next sub-section.

Pakistan-Based Learners

A group of 30 participants aged between 23 and 51 years (mean=32.66 sd=7.8) were selected from Pakistan. All of them were educated in Pakistan had obtained MA degrees from there. They had started listening to English as spoken by Pakistanis at an average age of 17.43 years (minimum=9, maximum=25, sd=4.3). The participants were selected from rural areas of southern Punjab in Pakistan. They had obtained education from public sector schools and colleges. Normally, in rural areas of Pakistan, students do not have opportunity to listen to English in schools. They are taught English through traditional grammar translation method. After they reach colleges, they have some opportunity to listen to Pakistani English. That is why the average age when the participants started listening to English is around 17 years. According to the self-reported statements, the participants speak English for an average of 1.23 hours daily (minimum=0, maximum=6, sd=1.5). They reported that they listen to English spoken by non-native speakers (Pakistanis) for an average of 0.63 hours per day (minimum=0, maximum=4, sd=.96). They never travelled to any English speaking country nor listened to English spoken by native speakers for a long time regularly. In other words, they only received input from Pakistanis. All of them speak central Saraiki as their L1.

UK-Based Learners

30 Pakistanis who were living in the UK (Essex) were selected as advanced learners. All of them are originally from the same areas from where the Pakistan-based participants were selected. All UK-based learners speak central Saraiki as their L1. They received input from both, Pakistani teachers and from native English speakers. Their ages range between 21 and 59 years (mean=33.26, sd=7.21) and their average length of residence (LOR) in the UK was 70.8 months (minimum=4, maximum=360, sd=80)[6]. All except two of the participants of the UK group in the current study had stayed in the UK for less than a year. The rest of them had stayed there for more than a year. They started listening to English spoken by Pakistanis at the average of 16.3 years (minimum=5, maximum=33, sd=8.15), and that by native English speakers after their arrival in the UK. Their age of arrival in the UK ranged between 19 and 36 years (mean=26.26, sd=4.58)[7]. 15 of them had either obtained an MA degree from the UK or were studying there, and 7 of them had got some diploma or certificate in their relative fields from the UK. According to their own statement, before coming to the UK, they listened to English as spoken by Pakistanis for less than an hour daily (mean=0.6, minimum=0, maximum=3, sd=.81) and they had never listened to native English spoken in Pakistan. They also speak the same dialect of Saraiki (central Saraiki) and had studied in the similar kind of educational institutions where the Pakistan-based participants studied. Extreme care was taken in selecting both groups of participants with a view that both groups are similar in all respects

6- Different studies consider different cut-off points for experienced learners, ranging from 6 months (Best & Tyler, 2007, p. 21) to as long as 42 years (Flege et al., 2003, p. 473). For example, it is 6-12 months in Best & Tyler (2007), Flege and Liu (2001) and Flege and Fletcher (1992), 1.3 years in Flege (1987), 5.1 years in Flege (1993) and 3.1 years in Guion et al. (2000).

7- Previous research shows that age of arrival has a vital role in acquisition of an L2.

except for one variable i.e. input. In this way, any improvement noticed in the UK participants will be ascribed to the input that they are receiving in England.

All participants of the UK group were living in Essex at the time of experiment. According to their own statements, they listen to English spoken by native speakers for an average of 5.50 hours daily (minimum=1, maximum=14, sd=3.59), and speak English for an average of 5.90 hours daily (minimum=2, maximum=12, sd=2.84). According to Schmid et al. (2014, p. 130) advanced L2 learners should be highly motivated, competent and have ample opportunity to receive native input. The main group of this study, i.e. UK-based learners fulfil these requirements. They are highly motivated because most of them who are living in the UK for long time are doing such jobs at shops which involve interaction with native speakers. Their job demands competence in speaking and listening English. A few of the participants who are studying in the UK and are here for less than a year time also have to interact with their teachers and peers in the classes. Besides, they have to go back to Pakistan and find job on the basis of their learning from the UK. In Pakistan, speaking good English is considered a strong indicator of being highly educated (Rahman, 2003). Thus they have to acquire excellent spoken English skills. Thus, the main group of advanced L2 learners qualify for being declared advanced adult learners of English. The predictions of SLM are also about advanced learners.

Control Groups

Two groups of participants were also selected as control. The control groups were English and Saraiki monolinguals. Twenty-two English monolingual speakers living in the same area of England (i.e. Essex) where the UK-based Pakistani learners were living, were selected for participation in this project. Ten Saraiki monolinguals were also selected for the same purpose from the same areas of Pakistan where the Pakistan-based adult Pakistani learners of English were living. The UK-based Pakistani learners of English had also been living in the same area in Pakistan before migrating to England.

The Experiment

Before the main experiment, participants were requested to fill a questionnaire which provided us the information reported above. They were also requested to give written permission on consent forms to collect information from them, record their voices to use for research purposes anonymously. The participants willingly signed the consent forms.

In the main experiment, participants were asked to read from a paper two lists of words which, along with some distracters, carried English words 'peak, speak, teeth, steal, key, ski'. One list comprised of only isolated words and the other had the same words embedded in a carrier sentence. The carrier sentence was 'I say ....again.' The target words were embedded between the word 'say' and 'again' in the carrier sentence. Each of the target words was randomly repeated three times in each of the lists. In this way, six productions (three in isolated words and three in carrier sentences) of each target word were recorded. The purpose of recording the target words in isolation and in continuous speech was to see differences in the performance of learners in two different contexts. Previous studies show that sometimes learners perform differently in production of isolated words and in continuous speech (Birdsong, 2007). Thus, competence in production of isolated words is, although necessary, not a sufficient condition for perfect L2 acquisition. Rather, competence in continuous speech implies the same in isolated words but not vice versa (ibid).

Voice onset time (VOT) of English plosives [p t k p^h t^h k^h] were measured using Praat (Boersma & Weenink, 2012). It is advisable to test within-speaker consistency in such studies (Foulkes, et al., 2010). Therefore, a Cronbach's alpha reliability test for consistency was applied to the six repetitions of the VOT values[8]. The test shows an ideal consistency in six repetitions of voiceless stops by the UK participants and good consistency among the Pakistan group. The following table shows results of the reliability test.

Table 1:
Cronbach's alpha values for consistency in repetitions

Sounds	UK	Pakistan
[P^h]	0.91	0.79
[t^h]	0.95	0.77
[k^h]	0.93	0.64
[p]	0.91	0.63
[t]	0.89	0.77
[k]	0.91	0.63

The above result shows that the Cronbach's alpha value in all cases is between 0.6 and 0.8 which is considered a cut off point of reliability for such research studies (Larson-Hall, 2010; Scholfield, 1995).

As pointed out earlier, in order to gain an idea of the average VOT of native speakers, VOTs for English plosives spoken by 22 male native speakers of English, who were from Essex, were also calculated. The native English speakers were requested to produce the words carrying the target voiceless stops (stimuli) which had already been used as the stimuli for recording the L2 learners. The average age of the native English participants was 42.36 years (minimum=18, maximum=65, sd=17.84). The VOT values of these participants will be used as a yardstick for comparison with the UK-based learners of English. The VOT of Saraiki stops were also measured from the recordings of the Saraiki monolinguals for comparison. (See appendix for the list of Saraiki words used as stimuli for getting VOT of the monolingual speakers of Saraiki). The VOT of English and Saraiki were also measured in words and sentences.

The VOT was calculated from the burst of stop to the onset of first complete vibration of vocal folds, as suggested by Cho & Ladefoged (1999). This pattern of measurement was strictly followed for all sounds because phoneticians define various methods for measuring VOT, but there is a consensus among them that it is vital to use a method consistently (Foulkes et al., 2010). The VOT values in words obtained from all participants were not significantly different from those obtained in sentences ($p>.1$). It is a standard practice to merge VOT values taken in different contexts if the values are not significantly different from each other (see e.g. Flege 1987). Therefore, two values (VOT in words and sentences) were merged and the average of six repetitions for each of the target sounds was taken for analysis. The following results are based on the means VOT values obtained in six repetitions by each of the participants.

8- Later analysis revealed that there was no significant difference between the VOTs obtained in sentences and isolated words. Therefore, the test was applied on six repetitions jointly and the repetitions were averaged for further analysis.

PRESENTATION

VOT values of two of the UK participants could not be measured due to technical reasons. Therefore, in the following discussion VOTs of 28 UK participants will be considered. Table 2 shows mean VOTs.

Table 2:
Mean VOT values

Allophones	Groups Sounds	Monolinguals		L2 learners	
		Saraiki[9] N=10	English N=22	Pakistan N=30	UK N=28
Aspirated	[p^h]	42.53	56.23	8.84	22.56
	[t^h]	60.33	80.00	19.31	25.60
	[k^h]	77.82	81.14	37.53	56.36
Unaspirated	[p]	11.12	10.16	10.10	17.70
	[t]	14.68	23.89	25.02	30.07
	[k]	29.90	29.43	27.78	37.61

tA three-way (2*3*4) repeated measures ANVOA, with aspiration and place of articulation as within-subject factors and grouping as between-subject factors, shows a strong interaction between these three factors (F=13.95, p<.0001). The interactions of place and aspiration (F=37.26, p<.0001), grouping and aspiration (F= 106.57, p<.0001) and grouping and place (F=7.56, p<.0001) are also significant. As far as the individual factors are concerned, differences at place of articulation (F= 295.86, p< .0001) and those for aspiration (F=353.94, p<.0001) contrasts are strongly significant. This means that the performance of the groups in production of the aspirated and unaspirated sounds is different. Similarly, the VOT of the participants in production of voiceless plosives of different places of articulation is significantly different. The overall increase in VOT is strongly linear (F=476.96, p<.0001), meaning that the VOT is increasing in stops from labial to coronal to dorsal place of articulation.

The current study is mainly concerned with acquisition of aspiration contrast by the learners. Although the over-all aspiration contrast is significant among the groups, but it is important to note that the monolinguals' VOT values were also included in the data. We know that both English and Saraiki monolinguals produce aspirated and unaspirated plosives with significantly different VOTs. It is quite possible that the group variance is significant only because of the monolinguals. Thus, for further confirmation, the aspiration contrast of only L2 learners was tested on three places (labial, coronal, dorsal) separately. The results show that aspiration contrast in the VOT values of the UK group is significant for dorsal stops only (t= 5.32, p< .0001), whereas for labial (t=1.82, p>.05) and coronal (p>.1) stops, aspiration contrast is non-significant for the UK group. This means VOTs of the aspirated and unaspirated allophones of labial and coronal stops are produced without any significant difference of VOT but the allophones of dorsal stop are produced with significantly different VOT ranges by the UK group of participants. Similarly, in the data obtained from the Pakistan group, the aspiration contrast is significant for coronal (t= -4.825, p< .0001) and dorsal (t= -5.574, p< .0001), but it is non-significant for the labial stop (p>.1). This means that the VOT of the

9- Saraiki has two coronal stops namely retroflex and dental. Initially, the VOT of both Saraiki coronals were taken but later analyses revealed that Saraiki learners of English related English /t/ with the L1 retroflex, therefore, the VOTs of Saraiki retroflex are given for the coronals. In the following discussion, only the IPA symbol [t] will be used for both Saraiki and English coronals.

Pakistan group of learners is similar in the aspirated and unaspirated labials, but different for coronal and dorsal stops.

The group variance ($F_{3,86}=38.79$, $p< .0001$) is also significant in the overall test applied on all groups together. However, post hoc Scheffe pair-wise comparisons show that the group variance is significant in all except one (UK-based learners & Saraiki monolinguals) pairs ($p<.0001$). This means that the Pakistan-based L2 learners' group is not only different from the UK group, but also from the Saraiki and English monolinguals, whereas the group variance between Saraiki monolinguals and the UK-based learners group is non-significant. Table 3 below shows the number of participants who are above, below or within the native VOT range. We follow Flege et al. (1999) in assuming a standard criterion for determining learners' accuracy in L2 sounds. They suggest considering all those cases as native-like which are within a range of two standard deviations on both sides of the mean values obtained by native speakers (ibid). Larson-Hall (2010) also considers this an effective method of determining native speaker range of performance. It is also commonly recommended that VOT should be studied in range rather than in fixed values (Kent & Read, 2002). Thus, following the criterion of Flege et al. (1999), M-2SD is the minimum and M+2SD the maximum point of native VOT range for English plosives (where M stands for the mean VOT of native speakers, and SD is standard deviation).

Summarising this result we can say that there are strong three-way and two-way interactions among three factors, namely place of articulation, aspiration and grouping. The increase in the VOT is linear from labial to coronal to dorsal. In the UK group, there is no significant difference between aspirated and unaspirated labial and coronal stops. However, VOT values of these participants are significantly different for aspirated and unaspirated dorsal stops. In the Pakistan-based group, the aspiration contrast for labial stops is non-significant, but it is significant for coronal and dorsal stops. Interestingly, the average VOT values of the unaspirated stops of English by the Pakistan-based group are bigger than those of the aspirated stops at the coronal position.

Table 3:

No. of participants within native VOT ranges of English stops

Sound	Native VOT range (in ms)	Relation to native range	PK (30)	UK (28)
[ph]	29.98 - 82.48	above	0	0
		within	0	8
		below	30	20
[th]	52.39 - 107.46	above	0	0
		within	0	2
		below	30	26
[kh]	54.53 - 105.96	above	0	0
		within	2	13
		below	28	15
[p]	4.03 - 16.29	above	6	13
		within	21	15
		below	3	0
[t]	11.60 -36.18	above	1	6
		within	29	22
		below	0	0
[k]	7.58 - 51. 28	above	0	4
		within	30	24
		below	0	0

We will analyze these results in the following section, but first a summary of the results is presented here. The whole result is summarized in the following figure[10].

(1)

UK group:	Mean VOT for [ph] = Mean VOT for [p]
	Mean VOT for [th] = Mean VOT for [t]
	Mean VOT for [kh] > Mean VOT for [k]
Pakistan group:	Mean VOT for [ph] = Mean VOT for [p]
	Mean VOT for [th] < Mean VOT for [t]
	Mean VOT for [kh] > Mean VOT for [k]

There is no significant difference between the VOTs of the UK participants and Saraiki monolinguals ($p>.1$). In all other pair-wise comparisons, the group variance is significant. We will discuss each of these points one by one in the following section.

ANALYSIS AND DISCUSSION

As already discussed, aspiration contrast is allophonic in English but phonemic in Saraiki. The Saraiki learners of English are already familiar with aspiration contrast in their L1. For acquisition of aspiration contrast in English what they need to do is to simply transfer the L1 phonemic contrast into the L2 allophonic contrast[11]. The results presented above show that Pakistan-based learners could not do this successfully. The VOT of the Pakistan-based learners for aspirated stops is significantly different from that of monolingual speakers of Saraiki and English ($p<.0001$). This means the VOT of English aspirated stops produced by the Pakistan-based learners is neither like the L1 nor like the L2. However, some improvement in the UK participants is apparent. This improvement is confirmed in the statistical tests in that, group variance between the Pakistan and UK participants is strongly significant ($p<.0001$). The post hoc comparisons show that the UK participants are significantly different from the Pakistan-based group ($p<.0001$) as well as from the native speakers of English ($p<.0001$), but they are not significantly different from the Saraiki monolinguals ($p>.1$). This means that, the UK participants did not produce English aspirated stops with native-like VOT but they produced English aspirated stops with their L1 VOT ranges. The results confirm that the UK participants maintain aspiration contrast in only velar stops. In other words, the UK group of learners produced aspirated and unaspirated velar stops of English like the corresponding L1 aspirated and unaspirated stops respectively.

Only these results support the idea of the SLM in that, the UK learners relate sounds of the L2 with the corresponding L1 velar stops at allophonic level. That is why, although the UK learners are different from the English native speakers, but on account of relating the allophones of English with the corresponding phonemes of their L1, they have acquired two separate representations for aspirated and non-aspirated allophones of English dorsal. However, this positive transfer from the L1 VOT ranges occurred with the UK learners only. The Pakistan group of learners did not transfer the L1 phonemic contrast into their L2 phonemic inventory. Thus, the improvement in the UK group can be ascribed to the input that

10- An equation symbol '=' means no significant difference and A>B indicates that A is greater than B and A<B indicates that A is smaller than B.

11- This contrasts with more commonly discussed scenario in the literature where L2 learners have to split an L1 allophonic contrast into a phonemic contrast in the L2 (see Eckman, Elreyes, & Iverson, 2003 for details).

the UK learners are receiving from native speakers of English. The UK group, after receiving input from native speakers, realize the existing difference between aspirated and unaspirated allophones of English plosives; they try to maintain this contrast in production but they could not acquire quite native-like VOT values (since the UK group of learners is significantly different from native speakers of English). Rather, they transfer the L1 phonemic contrast (since overall, the UK group of learners is not significantly different from the L1 monolinguals). This means that, the UK learners have acquired separate phonetic categories for the allophones of English velar stops on account of positive transfer from the L1 VOT values and these representations are deflected away from those of English monolinguals. This is the overall picture that emerges on the basis of statistical analysis of the participants as groups as a whole. If we consider the performance of individuals, we have a picture given in the following table for the participants with native-like VOT reproduced from Table 3.

Table 4:
Number of participants within native-like VOT ranges

Sounds	Pakistan	UK
[p^h]	0	8
[t^h]	0	2
[k^h]	2	13
[p]	21	15
[t]	29	22
[k]	30	24

By having a look at these results, we identify that performance of the UK learners is better than that of the Pakistan group in production of aspirated stops but on unaspirated stops, the Pakistan group seems apparently better than the UK group. Second, the performance of learners is significantly different on different places of articulation of plosives. For example, only 8 participants of the UK group have native-like VOTs for [p^h] and 2 are in native-like range in [t^h], but 13 are native-like in the VOT of [k^h].

We will analyze each of these points in turn. The first point is the group variance. Overall, the better performance of the UK group in production of the aspirated stops of English is understandable since the UK participants have direct interaction with native speakers of English, but better performance of the Pakistan group in production of unaspirated stops of English is unexpected. The reason for this is that the Pakistan-based learners equate both aspirated and non-aspirated stops of English with the unaspirated counterparts of their L1. The average VOT values of the unaspirated stops in Saraiki and English are 11.12 and 10.16 ms for [p] and 29.90 and 29.43 ms for [k], respectively (see table 2). The difference between these VOT values of the L1 and L2 unaspirated sounds is statistically non-significant ($p>.1$). However, the VOTs of unaspirated [t] and all three aspirated stops are significantly different in both languages. We take up the case of [p] and [k] here. (The discussion of unaspirated [t] and aspirated stops will follow).

The Pakistan-based learners have transferred their L1 unaspirated [p] and [k] to the L2 [p] & [k], respectively, and their performance is better because of this positive transfer. The question remains as to why they do not transfer the VOT of the aspirated stops of the L1 to the aspirated stops of the L2? (As we have seen the L2 VOT values of Pakistan-based learners for aspirated stops of English are significantly different from those of the aspirated stops of the L1 and L2 monolinguals). This is because normally English language teachers in universities and colleges of Pakistan also equate the aspirated stops of English with the unaspirated stops of their L1[12]. The question now arises why they equate both allophones of English stops with a single L1 (unaspirated) phoneme when they have two different phonemes in their L1 corresponding to two different allophones of English plosives?

The reason for this equivalence classification of both allophones of English with a single L1 phoneme is that the VOT values of aspirated stops in L1 and L2 are significantly different from each other. On the other hand, the VOT values of unaspirated English [p] and [k] and the VOTs of the corresponding stops in Saraiki are not significantly different from each other. Since the Pakistan-based learners of English do not have allophonic variance in aspiration for their L1, they can only identify the aspiration contrast at phonemic but not at allophonic level in English; they equate both allophones of English to a single L1 stop. For this equivalence they choose the L1 phoneme (i.e. unaspirated stop) which is closer to the either of two English allophones. Thus, the learners of English have motivation to assimilate both English allophones with the closest phonemes of their L1 (i.e. unaspirated stops). This equivalence is further strengthened by English orthography (which does not differentiate between aspirated and unaspirated stops whereas Urdu orthography maintains this difference) and particularly because of the learners' inaccessibility to native speech. Another important factor in this regard is Pakistani English which also neutralizes English aspiration contrast in favour of unaspirated allophone. All participants of this study initially acquired Pakistani English (PakE) in their schools and colleges. Therefore, they are under strong influence of PakE which does not maintain aspiration contrast at in English plosives at allophonic level (Rahman, 1990, 1991, Mahboob & Ahmar, 2004).

In the current context, one of the most important influencing factors is orthography. The orthographic representation of these sounds in the two languages is different. In Pakistan orthography, different graphemes are used for aspirated and unaspirated phonemes. Two letters are used for representation of aspirated stops in Pakistani languages spoken and written in southern Punjab (from where the participants were selected), a grapheme for the basic stop and another for the secondary articulation (i.e. aspiration, which is the grapheme also used for [h]). For example, the aspirated labial in Urdu (the national language) and Saraiki (the regional language and L1 of the participants), is written with graphemes which represent the sounds [p] and [h], while for the unaspirated labial sound only one letter is used which represents [p] sound. Since Urdu is the functional medium of instruction in Pakistan in public sector institutions from where the participants of this study have been selected, and the learners and their English language teachers do not have access to English native speech, the participants, as well as their teachers in Pakistani educational institutions, depend on orthography for pronunciation of English sounds. Thus, they do not differentiate between aspirated and unaspirated allophones of English on account of English orthography (which also does not differentiate between aspirated and unaspirated allophones of plosives). The influence of orthography/spelling on L2 sounds is already established in the literature (Hayes-Harb et al., 2010; LaCharite & Paradis, 2005, pp. 251-253).

12-This is a common practice among Pakistani learners to equate both allophones of English with the corresponding unaspirated phoneme of the L1 although most of the major languages of Pakistan like Urdu (Shmidt, 2007), Punjabi (Shackle, 2007), Sindhi (Khubchandani, 2007), Kashmiri (Koul, 2007) and Saraiki (Syed, 2013c) have aspiration contrast in voiceless stops at phonemic level. Owing to the similarity between Saraiki and other major Pakistani languages, the findings of this study are also valid for all Pakistani learners of English who speak Urdu, Punjabi, Kashmiri and Sindhi as L1s.

Another important reason for equating both allophones of English stops with one phoneme of the L1 is (as pointed out earlier), that the aspiration contrast in the L2 is allophonic but in the L1 it is phonemic. In L1 phonemic inventory of the learners, the change of aspiration implies a change in meaning. This would explain why Pakistani learners of English do not equate aspirated stops of the L2 with those of the L1; they simply transfer their L1 grammar which treats the two sounds as different phonemes.

Thus, two opposing trends in the performance of the Pakistan-based learners are apparent. They show very good performance in production of unaspirated stops of English but are very poor in production of aspirated stops. The reason for the former is positive transfer from the L1, and that of the latter is equivalence classification between unaspirated stops of the L1 and the English aspirated stops. This indicates that in the minds of Pakistan-based learners the allophones of English have single representation which equals to the L1 unaspirated stop. Hence, they produce unaspirated allophones of English [p] and [k] with native-like VOT as a free ride from this equivalence classification, resulting in native-like VOT for unaspirated stops of English but the same action for English aspirated stops results in poor production.

Since the UK-participants performed significantly differently, a question arises that why the UK participants do not do the same (i.e. transfer the VOT values of L1 unaspirated [p] and [k]) and obtain the same accuracy level in these sounds as obtained by the Pakistan group of learners). Had they done the same, they would have an equal number of participants in the native VOT range in production of unaspirated stops as the Pakistan group. But the results show that, whereas 21 and 30 participants of the Pakistan group of learners are in the native VOT range in production of English unaspirated [p] and [k] respectively, the UK group has 15 and 24 participants in the native range of VOT for these sounds (see table 4). Since the variance between the two groups is significant, it means this difference is meaningful.

This can be understood by having a look at the nature of errors made by the UK participants in acquisition of VOT of unaspirated stops (see table 3 above). The nature of errors by the UK group is of developmental type. The UK learners, after getting regular and abundant input from native speakers, realize the aspiration contrast in English. They increase quantity of aspiration in production of the aspirated phonemes. The problem is that, at this stage, they not only increase the quantity of aspiration in word-initial stops, but they also increase aspiration in the stops of English which should be produced without aspiration. This is reflected in the fact that there are more UK (rather than Pakistan) participants who produced English unaspirated [p k] with a VOT bigger than the native range. Table 3 shows that in the UK group, 6 participants produced [p] and 4 produced [k] with VOTs bigger than the native VOT ranges whereas this number is one and zero, respectively, for the Pakistan group. A correlation test shows medium size ($r=.46$, $p<.05$) positive correlation between VOT of aspirated and unaspirated stops of the UK group confirming that the UK group increase their VOT values for both aspirated and unaspirated stops of English. This is a developmental error resulting from hyper-correction as a result of which the UK learners apply a change to aspirated and unaspirated contexts resulting in the apparently poor performance on unaspirated stops of English. Thus, if we compare the number of participants in both groups who produced English unaspirated stops with VOTs only below the native range, the UK group does not seem to have performed poorer than the Pakistan group. Thus the UK group, that shows learning, has undergone some negative effects from this in contrast to the Pakistan group who only transfer their L1 grammar.

The second point raised above is variation in the performance of the learners in terms of place of articulation. Place of articulation is an important factor in acquisition of plosives. The statistical analyses show that variance in place of articulation is strongly significant. The interaction of place of articulation with grouping and aspiration is also significant, which means that different groups perform differently on different places of articulation. As discussed earlier, VOT values of stops in the L1 and L2 are not significantly different for unaspirated labial and dorsal stops. However, this difference is significant for unaspirated coronal stops and for all three aspirated stops. The case of the unaspirated coronal stop [t] will be discussed separately, for the moment, we keep in focus only aspirated labial and dorsal stops with reference to place of articulation.

The number of participants of the UK group in the native VOT range in production of aspirated stops also confirms a specific directionality of learning. Both the UK and Pakistan-based learners have significantly higher VOT values for the aspirated dorsal than for the unaspirated dorsal, which means both groups of learners have developed two different representations for the two allophones of English dorsal stop. Thus, it is established that maximum learning among the learners is in production of dorsal stops. In order to understand the reason for this, we need to compare the nature of laryngeal contrast in the L1 and L2. In the L1 and L2, labial and dorsal aspirated stops have the same phonological features. The difficulty for Pakistani learners in acquisition of aspirated stops of English arises because they equate aspirated stops of English with unaspirated stops of their L1 at initial stage of learning due to the specific reasons discussed above. In other words, they already produce unaspirated stops accurately as a result of positive transfer from the L1. Therefore, for accurate production of English stops, the learners have to increase quantity of aspiration or VOT in their production of English aspirated stops only. It has already been established that dorsal place is by default more amenable to increase in VOT with two reasons suggested for this. According to the first view, a short distance between the place of articulation and vocal folds yields a bigger VOT (Lisker & Abramson, 1964). The second view is that the wider area of contact between articulators gives a bigger VOT (Stevens et al., 1986). Dorsal place, compared with coronal and labial, lies closer to the vocal folds and has a wider contact area between the active and passive articulators. Thus, from both angles, dorsal stops are more amenable to bigger VOT values. That is why it has been observed that dorsal stops have comparatively greater VOTs in many languages of the world (Cho & Ladefoged, 1999; Kent & Read, 2002). In the present context, since the learners' target is to acquire greater VOT values for English aspirated stops. Thus, they acquired English dorsal stops with relative ease. This leads us to hypothesize that those Pakistani L2 learners of English who have only aspirated phonemes in their L1 (like speakers of Eastern Balochi and Persian spoken in Balochistan province of Pakistan) are expected to acquire English unaspirated dorsal stops later than labial and coronal unaspirated stops. The reason for this is that at initial stage of learning, they following the dominant trend in Pakistan, will equate unaspirated stops of English with their L1 aspirated phonemes and the task in front of them would be to develop a separate category of stops with short-lag VOT. Vis-a-vis velars, coronal and labial stops are more amenable to short-lag VOT; therefore these will be acquired before the dorsal stops of English. Previous studies (Lopez, 2012; Major, 1987) have come up with the same results.

The variable performance on dorsals vs. labials supports the idea that markedness, rather than just L1 transfer, also plays a role in L2 phonological acquisition (Eckman, 1977, 1991). If the effect of L1 were greater than that of universal markedness, the participants of this study would have performed equally in production of labial and dorsal aspirated stops. The results are rather different dorsal and labial places, which shows that markedness has a very strong role in phonological acquisition. The role of markedness in L2 acquisition is also attested in other studies (Broselow & Finer, 1991). However, SLM does not explicitly discuss it.

We now consider the case of coronal stops. The following observations on the results show that coronal stops are poorly acquired by the participants of this study:

a) Only 2 of the UK-based learners are in the native VOT range of unaspirated coronals against 13 and 8 for dorsal and labial respectively (see table 4).
b) The learners of both groups have bigger VOT values for the unaspirated coronal stops than for aspirated coronal stops of English (see table 2).

A comparison of the coronal stops in the L1 and L2 of the learners may help identify the reason for this. As discussed above, the difference between labial and dorsal stops of the L1 and L2 is only that of quantity of aspiration. But the case of coronal stops is different. English coronal stops are alveolar, but corresponding to this, Pakistani languages including the L1 of the participants (i.e. Saraiki) have dental and retroflex stops. Thus, vis-à-vis English coronal stops which are [+anterior, -distributed], the L1 of the learners has either dental stops with [+anterior, +distributed] or retroflex ones with [-anterior, -distributed] features[13]. These coronal stops exist in the phonemic inventory of all major Pakistani languages. Thus, Pakistani learners can equate English alveolar stops either with their L1 dental stops or retroflex stops. It has been observed that Pakistani learners of English produce English dental fricatives as dental stops (Rahman, 1991; Syed, 2013b). This means that their dental place is occupied by English dental fricatives, and only retroflex is left for English coronal stops. Thus, English coronal stops which are alveolar are substituted with L1 retroflex stops in Pakistani English (Rahman, 1991). Further analysis will reveal that the participants have produced English /t/ not only with an inaccurate VOT, but they also produced it as retroflex. This scenario is the same as we see in acquisition of English /s/ by Dutch learners. Dutch /s/ is produced with the tongue slightly curled back towards alveo-palatal zone (Quené, Orr & Van Leeuwen, 2017).

Another significant factor pointed out above is that the learners have greater VOT for [t] in ‘st’ clusters in the word 'steal' but smaller VOT in word-initial stop in the word 'teeth'. In order to understand the reason for this we need to understand the constraints on consonant clusters in Saraiki (the L1). The L1 of the participants does not allow ‘s+stop’ clusters syllable-initially. S+stop clusters (‘sp’ ‘st’ and ‘sk’) pose difficulty to Pakistani learners of English because major Pakistani languages (excluding only Pashto but including Saraiki) do not allow s+stop clusters. Word-initially, Pakistani learners produce English /t/ as retroflex (Mahboob & Ahmar, 2004; Rahman, 1990, 1991). But in production of ‘st’ clusters, they cannot curl back their tongue to produce a retroflex sound soon after [s] because [s] itself being [+anterior] requires an opposite (forward) position of the tongue tip. Thus, the ‘st’ cluster, which is in itself already difficult for the learners on account of being illegitimate cluster in their L1, becomes articulatorily further complicated when [t] (an intended retroflex in the L2 phonemic inventory of the learners) demands a backward tongue movement and [s] demands a forward movement of tip of the tongue. In the face of these difficulties, the learners produce English [t] in ‘st’ clusters with no [t]-retroflexion but rather like an alveolar [t]. When the learners move their tongue forward to realize a [+anterior] [s], it remains in the same position for production

13- The terms used in feature model by Clements and Hume (1995) are adopted in this description.

of the following [t] in the words like 'steal'. However, when these learners produce English [t] word/syllable initially, they produce it as retroflex because they do not have the constraint operative in this context which is operative in case of 'st' cluster. Thus, these learners produce English [t] as retroflex word/syllable initially in the contexts where the consonant occurs singly, but they produce it as alveolar in taut syllabic ‘st’ clusters word/syllable initially. The retroflex coronal stops have a smaller VOT by default compared with the alveolar (non-retroflex) [t] (Steriade, 2001). It is because of this articulatory reason that the VOT of the participants is relatively smaller in word-initial /t/ than in ‘st’ clusters. This shift from retroflex to non-retroflex (like alveolar) coronal by the learners cannot be ascribed to learning. It is rather a positive unintended outcome of the difficulty in articulation of an ‘s+retroflex t’ cluster. The influence of such articulatory constraints on acquisition has not been discussed in the SLM.

To confirm the assumption that the participants of this study have produced English /t/ in the word 'teach' with retroflexion and that in the word 'steal' without it, F3 of the productions of the participants were studied. If a consonant is produced as a retroflex, F3 of the adjacent vowel is lowered (Hamann, 2003; Ladefoged and Maddieson, 1996). Thus, F3 of vowels in words ‘teeth’ and ‘steal’ were taken in two phases, immediately after the burst for [t] when the following vowel is under the influence of the preceding stop and in the mid of the vowel when the formants have neutralized the effect of the adjacent consonant and are quite constant and stable. Since there were six repetitions (three in words and three in sentences), the Cronbach's alpha reliability test was also applied on the repetitions to determine consistency in productions. Table 5 shows the reliability coefficients. The reliability test results show that there is an excellent consistency (>80%) in the repetitions.

Table 5:
Consistency among the repetitions in F3

Group	F3 in the burst phase	Middle of F3 vowel
UK	0.804	0.914
Pakistan	0.836	0.825

The average height of the F3 taken in the mid vowel was deducted from the F3 taken immediately after the burst phase. If the F3 is lowered immediately after the burst, then the result should be in negative digits and if it is not lowered then it should be in positive digits or in zero. Table 6 shows mean difference taken by deducting mean height of the F3 in the middle of the vowel from mean F3 in the burst phase of /t/ in the words 'teeth, and 'steal'

Table 6:
Mean F3 raising/ lowering

ontext	Group	Minimum	Maximum	Mean	Std. Deviation
Steal	UK	-357.00	669.33	125.51	238.58
Teeth		-479.00	310.33	-17.59	224.56
Steal	Pakistan	-257.33	344.67	49.72	178.82
Teeth		-486.67	166.67	-99.12	154.02

Table 6 shows that overall, there is F3 lowering in production of 'teeth' but not in production of 'steal' which confirms that in production of the former (but not the latter), the target coronal was produced with retroflexion. A repeated measures analysis of variance shows that effect of context ('teeth' and 'steal') is significant ($F_{1,53}=17.53$, $p<.001$) but group difference is non-significant ($F_{1,53}=3.02$, $p=.09$). The interaction of group and context is also non-significant ($p>.1$). It means overall, both groups have produced /t/ in the word 'teeth' with equal retroflexion and both groups have produced English /t/ in the word ‘steal’ without retroflexion (i.e. alveolar). The following graph also reflects these differences.

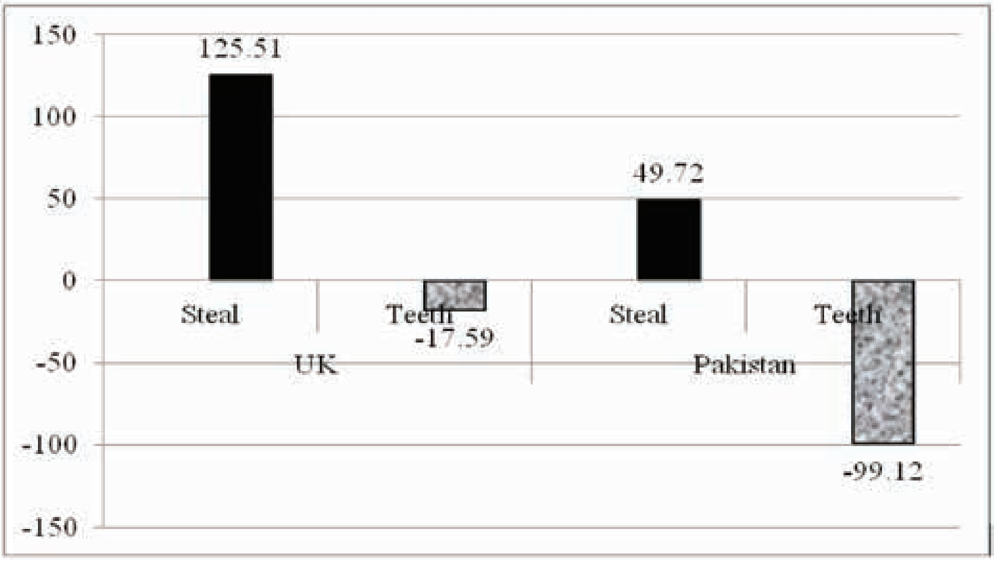

Figure 2: *F3 lowering/raising in the productions*

This clarifies the reason why Pakistanis produce /t/ in words starting with ‘st’ clusters ‘like in ‘steal’ with slightly longer VOTs compared with the contexts where the words start with single ‘t’ like in ‘teach’, etc. Now we see how these results fit in the SLM predictions. We shall also forward some suggestions to improve/amend the speech learning model.

A review of the speech learning model

For analyzing the findings of this study in the light of speech learning model, first we shall summarise the results. The results show that the participants have developed two separate representations for aspirated and unaspirated allophones of velar stop of English but they have equated both allophones of English labial with the L1 unaspirated labial stop. The performance of the participants is better in production of /t/ in the word 'steal' than in the word 'teeth'. In the word 'teeth' the target sound /t/ is produced as a retroflex whereas the same is produced as alveolar in the word 'steal' which is a result of articulatory constraints and negative transfer from the corresponding sounds of the L1.

The SLM predicts that learners relate sounds of L2 with the corresponding L1 sounds on position sensitive allophonic rather than on a more abstract phonemic level. The current study shows that participants relate allophones of English velar stop only with the corresponding L1 phonemes. In other words, only the results obtained from velar stop (but not those of the labial and coronal stops) of English support these predictions. These results also do not provide explicit support to the idea of the SLM that the mechanism which develops separate phonetic categories for new L2 sounds remains operative throughout life. The only example of somewhat accurate pronunciation of velar stops which the UK-participants managed to produce with two different VOT ranges is a result of positive transfer from the L1. Thus, only the idea of interaction between L1 and L2 is strongly verified by these results. According to one of the postulates of SLM, L1 and L2 phonemic inventories exist in a common phonological space which may result an interaction of the two phonemic inventories.

One important issue is related to the performance of the learners in acquisition of English coronals. From phonological point of view, labial and dorsal sounds are identical in both the L1 and L2 of the learners, but phonetically the aspirated allophones of English /p k/ have bigger VOT values than the corresponding L1 sounds. In other words, labial and dorsal stops are similar for the learners but the coronal stop is different (new), in that it not only differs from the closest L1 sound phonetically (in terms of VOT) but also phonologically (in terms of features). In this way, English coronals are new sounds for the participants. The SLM predicts that new L2 sounds are easier to learn than the similar ones, but the current study shows contradictory findings. Even, if we include English coronals in the list of similar sounds, the difference between the L1 and L2 coronals is bigger than that between dorsal and labial stops of the L1 and L2. SLM would predict more learning for the coronals than for the labial and dorsal stops of English for these learners. The results, however, do not conform to this prediction. It also points out a need to determine which particular yardstick L2 learners used to perceive the existing difference between two sounds. The current results demonstrate that Pakistani learners cannot perceive a difference between retroflex stops of their L1 and alveolar /t/ of English; that is why they develop an equivalence classification between these sounds utterly different sounds on word-initial position. As pointed out in section 1, SLM cannot determine what yardstick an L2 learner uses for perceiving L2 sounds. This shortcoming seems to have emerged as a solid hindrance in the way of accurate predictions in this L2 learning scenario. Thus, a revision is needed in the SLM for a clearer and more refined classification of new and similar sounds.

These results can be better understood if we differentiate between positive transfer and establishment of a new phonetic category for an L2 sound. For the SLM, development of a new phonetic category for an L2 sound does not necessarily mean that the developed representation is native-like. Recall that H6 of SLM assumes that learners may develop a phonetic representation that is ‘deflected away’ from the native category of target language. The VOTs of UK participants are not significantly different from those of the monolinguals of Saraiki in this study. This means that UK participants have the same VOT values for English dorsal stops as those in their L1 dorsal stops. This is an example of equivalence classification between L1 and L2 sounds. When Flege (1995) predicts that similar sounds are more difficult than new ones, he means that there is more probability of equivalence classification between two similar rather than between new sounds of the L1 and L2. This equivalence classification is confirmed in the case of dorsal stops in the current study. Thus, in line with the SLM, the dorsal stops of English which are similar to the L1 dorsal stops are equated with the L1 dorsal stops by the UK learners. And the improvement observed in production of velar stops of English by participants is a result of positive transfer. However, the equivalence between the L1 and L2 coronals and labials leads to negative transfer which yields worse results. In case of negative transfer, the equivalence classification reflects poor learning but in case of positive transfer the outcome seems better at the surface. The SLM cannot see a difference between positive and negative transfer and considers both as a means of blocking establishment of new category of an L2 sound. The SLM neither differentiates between negative and positive transfer from the L1 nor does it clearly define the concept of learning in such contexts.

Conclusively, two important findings need to be mentioned here. First, the interaction between markedness and learning is not suggested in the SLM. Although English labial and dorsal stops are similar sounds for participants of this study, they acquired proper VOT for aspirated dorsal stop more easily than for labial stop on account of naturalness of velar stops for bigger VOT. The SLM states that new sounds are acquired more easily than the similar ones, but it does not provide any directionality of learning between two sounds of the same type. It must be added to the SLM that among the similar sounds the more marked sounds are more difficult than the unmarked ones. Apart from markedness, other factors (such as orthography, articulatory constraints, etc.) may also affect acquisition of a new L2 sound as has been evidenced in this study. Particularly in the context of learning English as adult L2, the influence of orthography is very strong. These factors also need to be incorporated into a more comprehensive and revised version of speech learning model. Alternatively, as suggested by Lai (2009), no model of second language acquisition can develop across the board generalizations which encompass overall learning context of a particular community of learners. Adult learners exist and get influence from different factors. Various extra-linguistic factors also contribute in L2 acquisition. Thus, keeping in view predictions of different models and specific needs of a particular community, we need to develop individual models for different communities of learners which account for all linguistic and extra-linguistic factors operative influential in those particular communities. Different factors influencing L2 acquisition may contribute in L2 acquisition which may be identified by using latest research techniques.

Finally, as an aside, we refer to the production of English /t/ as retroflex in words like 'teach' and without retroflexion in words like 'steal. Better performance of the participants of this study in production of English /t/ in 'steal' is because of L1 and articulatory constraints. However, these findings also pose an utter challenge to the structure conformity hypothesis (SCH) which would claim that L2 learners acquire 'st' cluster after they have acquired 's' and 't' sounds accurately. As long as SLM is concerned, such factors have also not been discussed in the SLM explicitly. Therefore, findings of this study not only suggest a thorough revision in the SLM but may also be helpful in elaboration of other models of adult L2 learning.

CONCLUSION

The main objective of this project was to study acquisition of allophonic variance in plosives by adult Pakistan learners of English. The hypothesis for the study was (H1 of the SLM) that the learners would relate the allophones of the L2 (English) with the corresponding phonemes of the L1 (Saraiki) and, therefore, will acquire allophonic variance of English stops. The results show that the participants have acquired separate ranges of VOTs for aspirated and unaspirated allophones of English velar plosives only. But, they have equated aspirated and unaspirated allophones of English labials with the corresponding L1 unaspirated labial stop. Besides, they have produced /t/ in the word 'steal' with a VOT bigger than that in the word 'teeth'. In the word 'teeth' the target sound was produced as a retroflex. Thus the performance of participants was relatively better in production of /t/ in the word 'steal' which is due to L1 transfer and articulatory constraints.

The overall findings partially confirm the prediction (of SLM) that L1 and L2 phonemic inventories exist in the same phonological space and interact with each other. However, the predictions that new sounds are easier than the similar ones for L2 learners could not be verified in this study. The results also suggest that the SLM does not differentiate between negative and positive transfer. It also does not account for effect of markedness and the role of articulatory constraints in L2 acquisition. These issues need to be addressed in a revised version of the speech learning model.

APPENDIX

Stimuli for VOT of Saraiki Stops

Word	Meaning	Target Sound
[pi:la]	'yellow'	[p]
[ti:la]	'straw'	[t]
[tʰi:]	'be'	[tʰ]
[phi:ta]	'wheel'	[pʰ]
[ti:p]	'pomp'	[t]
[tʰi:k]	'right'	[tʰ]
[ki:ta]	'done'	[k]
[kʰi:sa]	'pocket'	[kʰ]

REFERENCES

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

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Timonium MD: York Press.

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

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis* 20, 3-49.

Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer.

Broselow, E., & Finer, D. (1991). Parameter setting in second language phonology and syntax. *Second Language Research*, 7(1), 35-59.

Browman, C., & Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, 18, 299-320.

Brown, C. A. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14(2), 136-193.

Brown, C. A. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 4-63). Malden Mass: Blackwell.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27(2), 207-229.

Clahsen, H., & Muysken, P. (1986). The availability of Universal Grammar to adult and child learners: A study of the acquisition of German word-order. *Second Language Research*, 2, 93-119.
Clahsen, H., & Muysken, P. (1989). The UG Paradox in L2 Acquisition. *Second Language Research*, 5, 1-29.
Clements, G. N., & Hume, E. V. (1995). The internal organization of speech sounds. In J. Goldsmith (Ed.), *A handbook of phonological theory* (pp. 245-306). Oxford: Blackwell.
Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris publications.
Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language learning*, 27(2), 315-330.
Eckman, F. R. (1991). The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition*, 13(01), 23-41.
Eckman, F. R., Elreyes, A., & Iverson, G. K. (2003). Some principles of second language phonology. *Second Language Research*, 19(3), 169-208.
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.
Flege, J. E. (1988). The development of skill in producing word-final English stops: Kinematic parameters. *Journal of the Acoustical Society of America*, 84(5), 1639-1652.
Flege, J. E. (1992a). The intelligibility of English vowels spoken by British and Dutch talkers. In R. D. Kent (Ed.), *Intelligibility in speech disorders thoery measurement and management* (pp. 157-232). Amsterdam: John Benjamins.
Flege, J. E. (1992b). Speech learning in a second language. In C. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 565-604). Timonium, MD: York press.
Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America*, 93(3), 1589-1608.
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. (1997). English vowel production by Dutch talkers: More evidence for the" similar" vs." new" distinction. In A. James & J. Leather (Eds.), *Second language speech: structure and process* (pp. 11-52). Berlin/New York: Mouton de Gruyter.
Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319-355). Berlin/New York: Mouton de Gruyter.
Flege, J. E. (2009). Give input a chance. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 175-190). Bristel: Multilingual Matters.
Flege, J. E., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America*, 83(2), 729-740.
Flege, J. E., & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91(1), 370-389.
Flege, J. E., & Liu, S. (2001). The effect of experience on adult's acquistion of a second language. *Studies in second language acquisition*, 23, 527-552.
Flege, J. E., Shirru, C., & Mackay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication* 40, 467-491.

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical Familiarity and English-language Experience Affect Japanese Adults' Perception of [ɹ] and [l]. *Journal of Acoustical Society of America*, 99(2), 1161-1173.
Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of memory and language*, 41(1), 78-104.
Foulkes, P., Docherty, G., & Jones, M. J. (2010). Analyzing Stop. In F. P, D. G. J & J. M. J (Eds.), *Sociophonetics: a student's guide* (pp. 58-71). London: Routledge.
Fox, R. A., & McGory, J. T. (2007). Second language acquisition of a regional dialect of American English by native Japanese speakers. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 117-134). Amsterdam: John Benjamins.
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.
Hao, Y.-C. (2014). *The Application of the Speech Learning Model to the L2 Acquisition of Mandarin Tones*. Paper presented at the 4th International Symposium on Tonal Aspects of Languages (TAL-2014), Nijmegen, The Netherlands.
Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *Journal of the Acoustical Society of America*, 110(1), 489-503.
Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the phonological forms of new words: effects of orthographic and auditory input. *Language & Speech*, 53(3), 367-381.
Honeybone, P. (2005). Diachronic evidence in segmental phonology: *The case of laryngeal specifications*. Berlin: Mouton de Gruyter.
Kato, K. (2005). *Second language (L2) segmental learning: Perception and production of L2 English by Japanese native speakers*. (PhD), University of Essex, Colchester.
Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech*: Singular San Diego, CA.
Khubchandani, L. M. (2007). Sindhi. In G. Cordona & D. Jain (Eds.), *The Indo Aryan Languages* (pp. 683-721). New York: Routledge.
Koul, N. O. (2007). Kashmiri. In G. Cordona & D. Jain (Eds.), *The Indo Aryan Language* (pp. 991-1051). New York: Routledge.
LaCharite, D., & Paradis, C. (2005). Category Preservation and Proximity versus Phonetic Approximation in Loanword Adaptation. *Linguistic Inquiry*, 36(2), 223–258.
Ladefoged, P. (1990). Some effects on the IPA. Journal of Phonetics, 18, 335-346.
Lado, R. (1957). *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor: University of Michigan Press
Lai, Y. (2009). Asymmetry in Mandarin affricate perception by learners of Mandarin Chinese. *Language and Cognitive Processes*, 24(7-8), 1265-1285.
Larson-Hall, J. (2004). Predicting perceptual success with segments: a test of Japanese speakers of Russian. Second Language Research, 20(1), 33-76.
Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.
Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.
Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, 20(3), 384-422.
Lopez, V. G. (2012). Spanish and English word-initial voiceless stop production in code-switched vs. monolingual structures. *Second Language Research*, 28(2), 243-263.

Major, R. C. (1987). English voiceless stop production by speakers of Brazilian Portugese. *Journal of Phonetics*, 15, 197-202.
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* 198-210.
Patkowski, M. S. (1990). Age and Accent in a Second Language: A Reply to James Emil Flege *Applied Linguistics*, 11(1), 73-89.
Penfield, W., & Roberts, L. (1959). Evidence from cortical mapping. *Speech and Brain Mechanisms*. 119-137.
Quené, H., Orr, R., & van Leeuwen, D. (2017). Phonetic similarity of /s/ in native and second language: Individual differences in learning curves. *The Journal of the Acoustical Society of America,* 142(6), EL519-EL524.
Rahman, T. (1991). Pakistani English: some phonological and phonetic features. *World Englishes*, 10(1), 83-95.
Rahman, T. (2003). *Language Education and Culture*. Karachi (Pakistan): Oxford University Press.
Schmid, M. S., Gilbers, S., & Nota, A. (2014). Ultimate attainment in late second language acquisition: Phonetic and grammatical challenges in advanced Dutch–English bilingualism. *Second Language Research*, 30(2), 129-157.
Schmidt, A. M. (1996). Cross-language identification of consonants. Part 1. Korean perception of English. *Journal of the Acoustical Society of America*, 99(5), 3201-3211.
Schmidt, R. L. (2007). Urdu. In G. C. D. Jain (Ed.), *The Indo-Aryan Languages*. New York: Routledge.
Scholfield, P. (1995). Quantifying language: *A researcher's and teacher's guide to gathering language data and reducing it to figures*. Clevendon/Philadelphia: Multilingual Matters Ltd.
Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. Cambridge: Newbury House
Shackle, C. (2007). Punjabi. In G. Cordona & D. Jain (Eds.), *The Indo Aryan Language*, 561-621.
Steriade, D. (2001). Directional asymmetries in place assimilation. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology*, 219-250.
Stevens, K. N., Keyser, S. J., & Kawasaki, H. (1986). Toward a phonetic and phonological theory of redundant features. Joseph S. Perkell and Dennis H. Klatt (eds) Invariance and Variability in Speech Processes.
Syed, N. A. (2013a). Acquisition of a new sound without separate category formation. In B. Surany (Ed.), *Proceedings of the Second Central European Conference in Linguistics for Postgraduate Students*. Pilicscaba, Budapest: Peter Pazmany Catholic University Hungary.
Syed, N. A. (2013b). Acquisition of dental fricatives by adult Pakistani learners. N*ewcastle Working Papers in Linguistics* 19, 2, 64-73.
Syed, N. A. (2013c). Voice onset time for plosives in Saraiki: Implications for the acquisition of English aspiration contrast. *EFL Annual Research Journal*, 14, 1-15.
Syed, N. A. (2014). Why Pakistani learners cannot acquire voiced stops of English? *Journal of Research and Social Sciences*, 1(1), 19-30.
Wardhaugh, R. (1970). The contrastive analysis hypothesis. *Tesol Quarterly*, 4(2), 123-130.
Wieling, M., Veenstra, P., Adank, P., & Tiede, M. (2017). *Articulatory differences between L1 and L2 speakers of English*. Paper presented at the The 11th International Seminar on Speech Production, Tianjin, China.