Available on Gale & affiliated international databases

ProQuest Start here. AsiaNet PAKISTAN

Journal of Humanities & Social Sciences

University of Peshawar

JHSS XXI, No. 3, 2013 (December)

Voice Onset Time (VOT) for Voiceless Plosives in Pashto (L1) and English (L2)

Nasir Abbas Syed

Department of English Language & Literature, Lasbela University, Baloshistan

---

Abstract

This is a study of voicing onset time for Pashto (L1) and English (L2) plosives with focus on acquisition of English plosives by adult Pashtoon learners. VOT for Pashto and English plosives were measured in carrier sentences. The results show that the overall direction of increase in the VOT for plosives in Pashto and English is from labial to coronal to velar but Pashto retroflex [ʈ] does not accord with this directionality. No influence of adjacent vowels on VOT of the preceding stops is noticed. The learners equate aspirated and unaspirated allophones of English labial /p/ and coronal /t/ with the corresponding L1 sounds neutralizing the aspiration contrast in the English plosives. However, they have separate phonetic representations for the allophones [k k^h] of English velar stop /k/. The findings of this study pose a challenge for feature model (Brown 1998, 2000) which predicts that a new L2 feature cannot be acquired by adult L2 learners whereas the participants of this study have acquired the feature [spread glottis] by developing two separate phonetic representations for the two allophones of English velar stop.

Keywords: Aspiration, L2 acquisition, Feature Model, Pashto, VOT

Introduction

Since the introduction of the contrastive analysis hypothesis (Lado, 1957) research on second language acquisition (SLA) has been mostly focused on comparative analysis of the L1 and L2 of learners. Several different models have been presented which identify origin of errors of adult L2 learners in L1 grammar.

Feature model (Brown 1998, 2000) predicts that problems in acquisition of new L2 sounds originate in the feature geometry of L1. The current study is based on acoustic analyses of Pashto (L1) and English (L2) stops produced by Pashtoon learners of English. Voice onset time is the most commonly studied acoustic correlate of plosives. Therefore, the current study focuses on voice onset time.

Voice onset time commonly called VOT is, in the words of Docherty (1992, p. 13), a term coined by Lisker and Abramson (1964) which denotes 'the interval (in ms) between the release of a stop closure and the onset of voicing for a following voiced segment'. In the classical study of Lisker and Abramson (1964), plosives are divided into three major types on the basis of VOT. If the voicing for the following segment starts soon after the burst of stops, such stops are called stops with short-lag VOT. English [p t k] and Pashto [p ʈ t̪ k] are examples of stops with short-lag VOT. If the voicing of the following segment starts fairly long after the burst of the stops, such stops are called plosives with long-lag VOT. English [p^h t^h k^h] are examples of stops with long-lag VOT. If the voicing of a plosive starts before the burst, such a stop is called pre-voiced or truly voiced stops and the VOT is measured in negative values. The phenomenon is called pre-voicing. Voiced stops are prevoiced in languages like Spanish (Flege & Eefting, 1988), Saraiki (Syed, 2012), Dutch (Simon, 2009), Japanese (Nasukawa, 2010), Arabic (Flege & Port, 1981; Simon, 2011), etc. The voiced stops in Pashto are also pre-voiced[1]. Phonologically, stops with long-lag VOT are called aspirated stops and those with short-lag VOT are called unaspirated stops. The VOT for unaspirated stops is below 40 ms and that for aspirated stops is above 40 ms in the world languages. The feature [spread glottis] differentiates between aspirated and unaspirated stops. In English, aspiration contrast is allophonic in that the aspirated [p^h t^h k^h] and unaspirated [p t k] stops do not make minimal pairs. In some languages like Urdu, aspiration contrast is phonemic because the aspirated and unaspirated stops make minimal pairs in Urdu. There are no aspirated stops in the phonemic inventory of Pashto (Elfenbein, 1997). In other words, Pashto does not have aspiration contrast and the feature [spread glottis] is not active in it. The current study focuses on voicing onset time for plosives in Pashto (L1) and English (L2).

The remainder of this paper is divided into the following six sections. Section 1 is about theoretical background of the feature model (FM) providing a brief review of the existing literature on the FM. Section 2 provides research questions and section 3 is about the research methodology. The results will be presented in section 4 and analysed and discussed in section 5. The paper ends with conclusion in section 6.

---

[1] To the best of my knowledge, there is no published work on VOT for voiced stops in Pashto. In a separate unpublished study I have measured VOT for voiced stops of Pashto.

1. Theoretical Background: Feature Model

The feature model by Brown (1998, 2000) is one of the potential models of SLA. The FM claims that acquisition of a new L2 sound depends on the feature geometry of L1 of learners. According to the FM, if two new L2 sounds which do not exist in the L1 are distinguished by a feature which is already active in the L1, learners will perceive the contrast between such a pair of the L2 sounds, and if the particular feature required to distinguish between the two L2 sounds is not active in feature geometry of the L1, the new L2 sounds will be difficult to perceive and acquire for the learners. If the feature required to distinguish the new L2 phonemes is inactive in the L1, the L1 feature geometry moulds the L2 phonemes according to the corresponding features of the L1 and the learners perceive the L2 sound the same as the closest L1 sound. This results in negative transfer from the L1.

The feature model is based on the findings of Brown resulting from empirical studies with speakers of some East Asian languages like Chinese, Japanese and Korean. The English consonant contrast [l r] is non-existent in Japanese and Chinese. However, according to Brown (1998), the feature [coronal] which differentiates between English [l r] is active in Chinese but inactive in Japanese. Brown (1998) found that Chinese learners could perceive the difference between [l] and [r] but Japanese learners could not perceive the same contrast accurately.

English [f v] and [b v] contrasts do not exist in Japanese but the relevant features [voice] and [continuant] which are required to differentiate between the consonants of the pairs [f v] and [b v] are active in Japanese. English [f] is [-voice] and [v] is [+voice]. Therefore, the feature [voice] is required to discriminate between [f] and [v]. There are some other pairs of sounds in Japanese which are also discriminated on the basis of the feature [voice]. This means the feature [voice] is active in Japanese. Similarly, the feature [continuant] which is required to differentiate between [b] and [v] is also active in Japanese. However, the feature [coronal] which is needed in the discrimination of English [l] from [r] is not active in Japanese. Brown (1998) studied the perception of English [b v], [f v] and [l r] contrasts by Japanese learners. The results show that the participants could only discriminate [v] from [f] and [b] but not [l] from [r].

Brown (2000) also studied the perception of English [p f], [b v], [f v], [s θ] and [l r] contrasts by Japanese, Chinese and Korean learners. English [s] is [-distributed] and [θ] is [+distributed]. Thus the feature required to discriminate between [s] and [θ] in English is [distributed]. The features required to discriminate between the members of the other pairs as discussed above, are [coronal], [continuant] and [voice]. These contrasts do not exist in Japanese, Korean and Mandarin Chinese. However, the features [continuant] and [voice] are active in Japanese, Korean and Mandarin Chinese and the feature [coronal] is active in Mandarin Chinese only. The feature [distributed] is not active in these languages. In this context, the feature model predicts that the pairs of English consonants [p f], [f v] and [b v] are easy to acquire for Japanese learners because the feature [voice] which discriminates between [f] and [v], and the feature [continuant] which discriminates between [p] and [f] and [b] and [v] are active in Japanese. But [s θ] and [l r] contrasts will be difficult for them to perceive because the features [distributed] and [coronal] are not active in their L1. Similarly, Chinese learners of English can easily perceive the difference between the members of the English contrasts [p f], [b v], [f v] and [l r] because the relevant features are active in their L1. Korean learners may acquire [p f], [b v] and [f v] contrasts owing to the same reason. However, the [s θ] pair of English consonants may be difficult for Korean and Chinese learners. The [l r] pair may also be difficult for Korean learners because of the L1 feature geometry. The results of the study by Brown (2000) were according to the predictions of the feature model. The performance of the Japanese and Korean participants was poor on [l r] and [s θ] contrasts, and that of the Chinese on [s θ] contrast only. All the participants performed excellently on all other contrasts. On the basis of these experiments Brown (2000) concluded that new L2 features cannot be acquired but new L2 sounds can be acquired if the relevant features are already active in the L1 of learners.

Larson-Hall (2004) points out two shortcoming of the FM; 1) that it gets most of its empirical support from Brown's own experiments, and 2) that it mostly focuses on the speakers of a specific group of languages; i.e. East Asian languages. The current study focuses on speakers of Pashto which is a language of a different family. The feature [spread glottis] which is required to discriminate between aspirated and unaspirated consonants of English is not active in Pashto. Accordingly, Pashtoon learners of English cannot perceive the difference between aspirated and unaspirated stops of English. Consequently, they may not acquire the aspiration contrast in English. The main aim of the current study is to test the prediction of the FM. According to our knowledge, there is no previous study on voice onset time for stops in Pashto. Therefore the study of VOT for Pashto voiceless stops will also make part of this experiment. The influence of vowels on the adjacent stops will also be studied. The study is mainly concerned to the following research questions.

2. Research questions

1. What is the voice onset time (VOT) for voiceless stops in Pashto?
2. Is there any influence of the adjacent vowel on the VOT?
3. Can Pashtoon learners acquire aspiration contrast of English which is non-existent in Pashto?

To address these questions we conducted a production test with a group of 12 Pashtoon learners of English. The details of the experiment and the participants are given in the following section.

3. Research Methodology

Twelve Pashtoon learners of English who were from Khyber Pakhtoonkhaw province of Pakistan, participated in this experiment. Eleven of them were recorded in Colchester (Essex) and one in London. All of them were university teachers in Pakistan doing PhD in England. The participants did not accept any reimbursement for their time. They rather participated in the experiment voluntarily.

The participants were asked to read words of English and Pashto (stimuli) in carrier sentences. The stimuli carried voiceless plosives. The carrier sentence for English was 'I say ……….again' and that for Pashto was 'da …..dei' (*This is …..').* The target sounds were recorded in the carrier sentences to maintain naturalness in the productions. Each of the target sounds was produced in the context of three vowels i.e. [i a u]. It was extremely difficult to find out Pashtoon monolinguals in England. Therefore, four of the twelve participants who produced English words were asked to produce the words of Pashto carrying the target sounds each three times. (See the list of the stimuli in the appendix). In this way, we obtained 36 productions (3 vowels*12 participants for English and 3 vowels*3 repetitions*4 participants for Pashto) of each of the target consonants. Each of the productions of the Pashto sounds was considered a case in the quantitative data analysis. In this way, both sets of English and Pashto plosives had equal number of recordings for analysis. For English, three aspirated [p^h, t^h, k^h] and three unaspirated stops [p t k] stops and for Pashto labial /p/, laminal coronal /t̪/, retroflex /ʈ/, and velar /k/ stops were the target sounds. Only voiceless sounds were studied in the experiment. The productions were recorded using M-Audio Track-II digital recorder and analyzed using Praat (Boersma & Weenink, 2012). Following the standard method commonly used in the literature, the voice onset time was measured from the burst of the stop to the beginning of first complete cycle of periodic vibration of the vocal folds (Foulkes, Docherty, & Jones, 2010). Although the measurement was based on the waveforms, the spectrograms were also considered for determining the burst and the onset of the periodic noise for the following vowel.

Before recording, the researcher conducted an interview with the participants. The interview carried questions regarding their age of arrival and length of residence in the UK, speaking listening habits, etc. The detail of the participants given in the following table is based on the information obtained in the interview.

Table 1: Detail of the participants

	Minimum	Maximum	Mean	Std. Deviation
Age (years)	27.00	38.00	32.75	3.82
Age of arrival in UK (years)	24.00	36.00	30.00	3.79
Length of residence in UK (months)	4.00	96.00	29.17	24.20
Speaking English hours/day	1.00	8.00	4.25	2.70
Listening English hours/day	1.00	10.00	5.25	2.60

The above table shows that the participants of this study were of an average age of 32.75 years who arrived in the United Kingdom at an average age of 30 years. According to their own statement, they listen to English for approximately 5.25 hours daily and speak it for an average of 4.25 hours daily. Their average length of residence in the UK was 29.17 months.

4. Results

The results of the experiment are presented in this section. The voice onset times for plosives in Pashto (L1) are presented in 4.1 and for English (L2) plosives are presented in 4.2.

4.1. VOT for Pashto stops

In a repeated measures analysis of variance (RM ANOVA), the effect of vowel on the preceding stops was found to be non-significant (p>.1) in Pashto. Therefore the repetitions obtained in the three vocalic contexts were averaged. The following table shows the average VOT for Pashto stops.

Table 2: VOT for voiceless plosives in Pashto

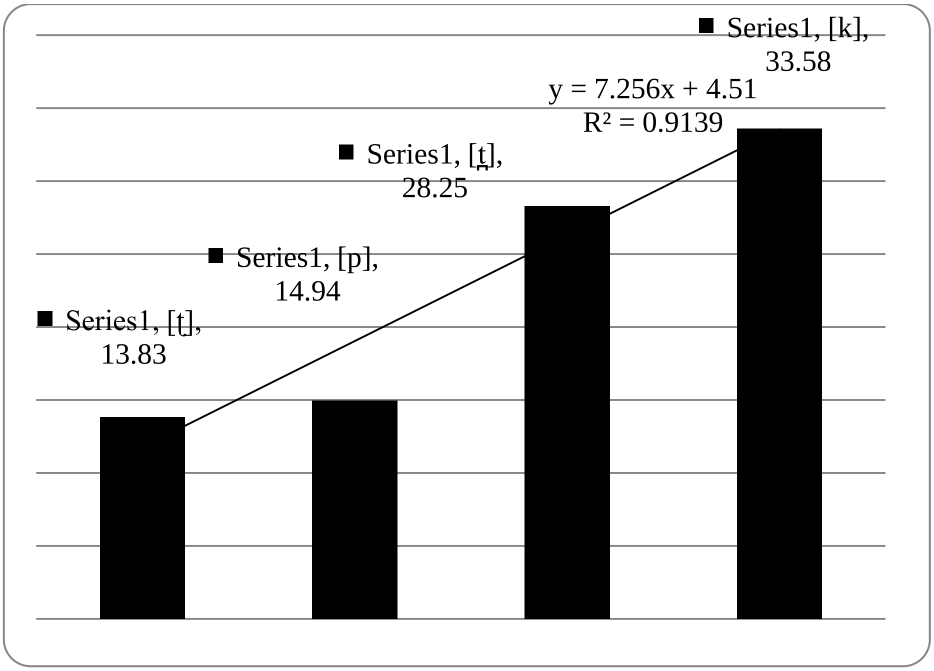
Sound	Minimum	Maximum	Mean	Std. Deviation
[p]	6.00	30.00	14.94	5.86
[t̪]	15.00	45.00	28.25	7.96
[ʈ]	6.00	30.00	13.83	6.58
[k]	18.00	55.00	33.58	10.09

The overall effect of place of articulation is signficant ($F_{3,11}$=44.82, p>.001)[2] but the effect of vowel on the adjacent consonant is non-significant (p>.1). This means the VOT for voiceless stops of Pashto at different places of articulation are significantly different from each other. However individual comparisons show that

[2] A non-parametric KS test confirms that the data are not significantly (p>.1) different from normal distribution and hence qualify for parametric analysis.

the VOT of labial and retroflex stops of Pashto are not significantly different from each other (p>.1). Retroflex sounds always have smaller VOTs than non-retroflex sounds (Steriade, 2001, pp. 224-225). Therefore, the findings of the current study are in line with the existing literature. If we exclude retroflex stops, the remaining plosives of Pashto have a linear increase of VOT from labial to coronal to velar which is reflected in the following figure.

Figure 1: Linear increase in the VOT for Pashto plosives



4.2. VOT for English stops by Pashtoon learners

The effect of vowel was found non-significant (p>.1) on the VOT of English stops produced by the participants. Therefore the repetitions were averaged. The following table shows average VOTs for English stops produced by the participants.

The overall effect of place (Wald Chi-square=43.182, p>.001)[3] and aspiration contrast (Wald Chi-square= 6.832, p=.009) is significant whereas the effect of the following vowel on the VOT of the preceding consonant is non-significant (p>.1).

[3] A non-parametric KS test confirms (p<.05) that the data do not qualify for parametric test. Therefore a non-parametric test was applied.

This means that the VOTs of the aspirated stops are significantly bigger than those of the unaspirated stops. The direction of increase in the VOT is from labial to coronal to velar in both aspirated and unaspirated stops. However, the adjacent vowels do not change the VOTs of the preceding stops.

Table 3: Average VOT of Pashtoon learners of English

Sound	Minimum	Maximum	Mean	Std. Deviation
[pʰ]	5.00	66.00	22.44	15.97
[p]	6.33	47.33	18.47	12.29
[tʰ]	12.67	98.67	35.64	22.88
[t]	17.67	61.33	33.17	11.84
[kʰ]	37.67	82.33	59.03	15.38
[k]	24.67	79.33	46.94	13.53

The most important finding in the above results is that the aspiration contrast is significant in the productions of the participants, which shows that the learners have developed two separate ranges of voice onset time for aspirated and unaspirated stops of English. We already know that the L1 of the learners has only unaspirated stops. The above table shows that the VOT of labial and coronal aspirated and unaspirated stops are very close to each other. Therefore, for further confirmation of the above analysis we compared the mean VOTs of the aspirated and unaspirated stops separately. The results show that the differences between the VOTs of the aspirated and unaspirated labial and coronal stops produced are non-significant ($p>.1$). Only the difference between the aspirated and unaspirated velar stops is significant ($Z= -2.907$, $p=.004$). This means that the overall significant difference between the VOT values of the aspirated and unaspirated stops is based on only the difference between the two allophones of velar stops.

This confirms that the learners have developed two separate representations for the allophones of velar stops of English although they have similar phonetic representations for aspirated and non-aspirated labial and coronal stops of English. In other words, there is some learning observed on velar stops but no significant learning for labial and coronal stops of English is observed in the performance of the participants. It raises a question whether the learners have transferred the L1 VOT values for some of the L2 stops or developed separate representations on account of learning. In other words we need to tease apart the effect of positive transfer from the L1 and development. The mean VOT values for Pashto and English plosives are compared to identify the effect of the L1 and development on the acquisition of the L2. The following table shows the results of the test.

Table 4: Comparison between the VOTs for Pashto and English plosives

Sounds	Z	Asymp. Sig. (2-tailed)
English [p^{h}] & Pashto [p]	-1.90	0.057
English [p] & Pashto [p]	-0.43	0.664
English [t^{h}] & Pashto [t̪]	-0.28	0.778
English [t] & Pashto [t̪]	-0.89	0.376
English [k^{h}] & Pashto [k]	-5.23	0.000
English [k] & Pashto [k]	-2.94	0.003
English [t^{h}] & Pashto [ʈ]	-5.42	0.000
English [t] & Pashto [ʈ]	-5.76	0.000

The above table confirms that there is no significant difference between the L1 labial and laminal coronal stops and the allophones of the L2 labial and coronal stops. This confirms that the participants have transferred the L1 VOT values for the aspirated and unaspirated allophones of labial /p/ and coronal /t/ stops of English. However, they have developed separate ranges of voice onset time for allophones of velar /k/ stop of English which are significantly different from the L1 velar stop. We shall further analyze and discuss these results in the following section.

5. Analysis and discussion

We can summarize the above results in the following points;

1. The effect of vowels on the voice onset time for stops is non-significant in Pashto and English.
2. The direction of increase of VOT for plosives in Pashto (L1) and English (L2) is from labial to coronal to velar but the retroflex sounds of Pashto do not accord to this pattern.
3. The learners equate aspirated and unaspirated labial /p/ and coronal /t/ of English with the corresponding L1 stops but develop two separate VOT ranges for the aspirated [k^{h}] and unaspirated [k] allophones of English velar stop /k/.

We will comment on these points one by one. The findings of this study are different from some of the previous studies which have found significant effect of the adjacent vowels on the acquisition of L2. Iverson et. al (2008), Johnson and Babel (2010) and Kluge et. al (2007) found strong effect of vowel on the adjacent consonants. Syed (2011) found significant effect of the adjacent vowels on the perception of English voiced alveo-palatal fricative by Pashtoon learners but the effect of vowel on the perception of other consonants of English was not significant at the alpha level of .05 in the same study. In the current study, the effect of the vocalic context is neutral on the plosives of the L1 and L2.

The direction of increase in the voicing onset time for plosives in Pashto (L1) and English (L2) is from labial to coronal to velar. This is in line with the previous findings. Cho and Ladefoged (1999) studied VOT for stops in 18 Indian languages and found the same direction of increase in the VOT in most of the languages. Lisker and Abramson (1964) also predict the same direction of increase of VOT for plosives. It has been observed that VOT of stops is inversely proportional to the distance between the vocal folds and the place of articulation. Long distance between the place of articulation of stops and the vocal folds yields smaller VOT and short distance between them yields bigger VOT (Stevens, Keyser, & Kawasaki, 1986). The reason for this is that the air coming from lungs is compressed between the vocal folds and the point of contact between active and passive articulators. The pressure is high if the area of the compressed air is smaller and low in a big area. A highly built pressure takes longer to normalize than a low-built pressure. The vocal folds start vibrating when the organs of speech come to their normal position. In this way, a longer distance between the place of articulation and the vocal folds yields a smaller VOT. That is why labial stops have small but velar stops have big VOTs.

Another point of view is that a wider contact area between active and passive articulators yields a big VOT (Lisker & Abramson, 1964). The reason for this is that if the contact area between the articulators is wide, it will take longer for the organs of speech to separate and normalize. On the other hand if the contact area between the articulators is smaller, it will take a short duration of time for the organs to separate and normalize. That is why velar stops having bigger contact area between the articulators have relatively bigger VOT and coronal stops having smaller contact area between active and passive articulators have smaller VOT in most languages of the world. From both points of view, velar stops are predicted to have the biggest VOTs. The findings of this study are also in accordance with the attested pattern.

The most important finding of this study is that the learners have developed separate phonetic representations for English aspirated and unaspirated allophones of velar stops. The mean VOT of /k/ in Pashto is 33.58 ms (see table 2) but for English [k] and [k^h] is 46.94 ms and 59.03 ms (see table 3), respectively. This means that the learners have not transferred the L1 VOT values for either of the allophones of English /k/ whereas they have transferred their L1 VOT values for both allophones of English labial /p/ and coronal /t/ neutralizing the aspiration contrast in the L2. The learners would have developed two categorically separate VOT ranges had they transferred the L1 VOT values for the unaspirated allophones of English plosives and developed separate VOT ranges for the aspirated allophones of English stops. But the results show a different picture. Whereas they equated both allophones of English labial and coronal stops /p t/ with the corresponding L1 stops, they developed two separate ranges of VOTs for the allophones of velar stops of English. This indicates real development. Previous research shows that most Pakistanis including Pashtoon learners neutralize the aspiration contrast in English plosives and produce both aspirated and unaspirated allophones of English stops as unaspirated (Mahboob & Ahmar, 2004; Rahman, 1990, 1991). When the Pashtoon learners realize the aspiration contrast in English on account of native input in the UK they increase the VOT of the aspirated stops of English. However, they over-generalize it and increase the VOT of unaspirated velar stops of English as well. Such overgeneralization is an expected outcome of hyper-correction which, in the current context, may be considered a developmental error (Major, 2001, 2008).

Table 3 shows that the average VOT for English voiceless aspirated labial [p^h] is 22.44 ms and unaspirated labial [p] is 18.47 ms whereas the average VOT for the corresponding Pashto stops is 14.94 ms. Although the difference between English [p^h] and Pashto [p] is non-significant (p=.057) the null hypothesis is rejected with a very narrow margin. Some linguists (Larson-Hall, 2010) recommend rejection of the null hypothesis at the level of .1 alpha value. Keeping this in view, we can claim that the VOT for Pashto [p] is significantly different from that of English [p^h] but not from English [p]. In other words learners have equated the unaspirated labial [p] of English with the corresponding Pashto [p] but developed a separate VOT value for the English aspirated labial [p^h]. Therefore, we consider the learners in the process of development and claim that they are likely to develop two separate VOT ranges for the allophones of all voiceless plosives of English.

Finally, we need to analyze the results of English coronal. We already know that most Pakistan-based learners equate English /t/ with their L1 retroflex sounds (Mahboob & Ahmar, 2004; Rahman, 1990, 1991). In the current study, we found that the UK-based Pashtoon learners have the average VOT values for the allophones of English /t/ (35.64 ms for the aspirated and 33.17 for unaspirated allophone), which are significantly different from the L1 retroflex [ʈ] (13.83 ms). However, their VOTs for English /t/ are not significantly different from the L1 laminal coronal [t̪] which is 28.25 ms. (see tables 2 and 3 above). This indicates that the performance of the UK-based Pashtoon learners is different from Pakistan-based

Pakistani learners in that, unlike the latter, the former did not transfer the VOT values of the L1 retroflex sound to the L2 /t/. These results confirm that the participants of this study are on the way to acquire accurate VOT ranges for English plosives and the direction of learning is from velar to labial to coronal.

This confirms that L2 learners acquire unmarked sounds before the marked ones (Archibald, 1998). As observed in most of the world languages, velar stops are the most unmarked for big VOT values. Pashto has unaspirated stops but not aspirated ones. Therefore the main task for the Pashtoon learners of English was to develop a new phonetic category for aspirated stops with VOTs significantly bigger than the unaspirated stops. The velars stops being most amenable for big VOT values are acquired first which confirms the role of markedness in L2 acquisition.

These findings pose a possible challenge for the feature model which claims that adult L2 learners cannot acquire a new feature in the L2 (Brown, 1998, 2000). We know that the feature [spread glottis] is not active in Pashto. But the participants of this study have developed two different representations for English [k] and [k^h] in their L2 phonemic inventory and are likely to acquire English /p t/ within the due course of time. According to Flege (1995), acquisition of separate phonetic categories for two L2 sounds means acquisition of the L2 sounds although the new representations are deflected away from the phonetic representations of monolinguals of the L2. Thus, the participants of this study have learnt aspiration contrast in English velar stops. These findings contradict the claims of the FM and point out a need for revision in the model.

The learning observed in the performance of the participants may be ascribed to the prevalent linguistic situation in Pakistan. Urdu being a national language of Pakistan is a lingua franca in the country. Almost all educated Pakistanis know Urdu very well. Urdu has aspiration contrast at phonemic level. This is yet to be determined whether these learners acquired aspiration contrast in the velar stops of English under the influence of already acquired aspiration contrast in Urdu or because of the input they are receiving from native speakers of English in the UK or both. It is also noteworthy that if they had transferred Urdu aspiration contrast to English, they would have performed equally well on all sounds. These questions need to be addressed in the future research. Anyway the findings of this study pose a possible challenge for the feature model which claims that a new L2 feature may not be acquired.

6. Conclusion

The current study was focused on the VOT of Pashto and English plosives. The results show that vowels do not have any effect on the VOTs of the adjacent stops. The direction of increase in the VOT of Pashto (L1) and English (L2) is from labial to coronal to velar with the exception of Pashto voiceless retroflex consonants. The findings demonstrate that, contrary to the claims of the feature model, the Pashtoon learners have acquired aspiration contrast in English velar stops. This points out a need for revision in the feature model. However, it must also be noted that the study was focused on the aspiration contrast in English which is only allophonic (not phonemic). For further confirmation of these findings, acquisition of aspiration contrast may be studied at phonemic level with a relatively larger sample.

Appendix

A. The following words of English each written once on a paper were read by 12 participants in the carrier sentence 'I say ……..again.' Peak, speak, pool, spoon, park, spark, teeth, steal, tall, stall, tool, stool, key, ski, car, scarf, coup, scooter.

B. The following words of Pashto each written three times on a paper were read by 4 participants in the carrier sentence, 'da ………..dei' "*this is ……….'. pa*ʈ (lying), peeʈa (chaff), pokh (paved, cooked), ʈeera (sword), ʈala (lock), ʈoot (melburry), ʈikala (loaf), ʈal (wood-merchant shop), tokha (cough), keegee[4] (be, happen), kargha (crow), kooza (water-pot).

[4] The Pashto word 'keegee' was produced in the sentence 'da say keegee'.

References

Archibald, J. (1998). *Second language phonology* (Vol. 17). Amsterdam/Philadelphia: John Benjamins Publishing Company.

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

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). Maldan Mass: Blackwell.

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

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris publications.

Elfenbein, J. (1997). Pashto Phonology. *Phonologies of Asia and Africa 2*(2), 733-759.

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., & 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., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language & Speech, 24*, 125-146.

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): Routledge.

Iverson, P., Ekanayake, D., Hamann, S., Sennema, A., & Evans, B. G. (2008). Category and perceptual interference in second-language phoneme learning: An examination of English /w/-/v/ learning by Sinhala, German, and Dutch speakers. *Journal of experimental psychology: human perception and performance, 34*(5), 1305.

Johnson, K., & Babel, M. (2010). On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics, 38*(1), 127-136.

Kluge, D. C., Rauber, A. S., Reis, M. S., & Bion, R. A. H. (2007). *The relationship between the perception and production of English nasal codas by Brazilian learners of English.* Paper presented at the Interspeech, Antwerp.

Lado, R. (1957). *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor: University of Michigan Press

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.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word, 20*(3), 384-422.

Mahboob, A., & Ahmar, N. H. (2004). Pakistani English: Phonology. In E. W. Schneider (Ed.), *A Handbook of Varieties of English: A Multimedia Reference Tool* (pp. 1003-1015). Berlin: Mouton de Gruyter.

Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah NJ: Lawrence Erlbaum Associates.

Major, R. C. (2008). Transfer in second language phonology: A review. In E. Hansen & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 63-94). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Nasukawa, K. (2010). *Place-Dependent VOT in L2 Acquisition.* Paper presented at the The 2008 SLR Forum, Somerville.

Rahman, T. (1990). *Pakistani English: The linguistic description of a non-native variety of English*. Islamabad: National Institute of Pakistan Studies, Quaid-i-Azam University.

Rahman, T. (1991). Pakistani English: some phonological and phonetic features. *World Englishes, 10*(1), 83-95.

Simon, E. (2009). Acquiring a new second language contrast: an analysis of the English laryngeal system of native speakers of Dutch. *Second Language Research, 25*(3), 377-408.

Simon, E. (2011). Laryngeal stop systems in contact: Connecting present-day acquisition findings and historical contact hypotheses. *Diachronica, 28*(2), 225-254.

Steriade, D. (2001). Directional asymmetries in place assimilation. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219-250). London: Academic Press.

Stevens, K. N., Keyser, S. J., & Kawasaki, H. (1986). Toward a Phonetic and Phonological Theory of Redundant Features. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 426-463). Hillsdale, N.J: Lawrence Earlbaum.

Syed, N. A. (2011). Perception and production of consonants of English by Pashto speakers. *The Journal of Humanities and Social Sciences, University of Peshawar, XIX*(1), 119-146.

Syed, N. A. (2012). Perception and production of English [d] by Pakistani L2 learners. *Essex Graduate Students Papers in Language and Linguistics, University of Essex, 13*, 134-157.