Available on Gale & affiliated international databases

University of Peshawar

JHSS XIX, No. 1, 2011

Journal of Humanities & Social Sciences

Perception & Production of Consonants of English by Pashto Speakers

Nasir Abbas Syed

Department of English Language and Linguistics
Lasbela University of Agriculture, Water and Marine Sciences, Uthal, Balochistan, Pakistan
(Currently PhD Scholar, University of Essex, UK)

Abstract

This paper deals with the perception and production of L2 consonants by Pashtoon learners/speakers of English. Phonetic, phonological, perceptual and social factors are considered to account for the problems that adult Pashtoon learners face in the acquisition of new and similar consonants of English. The findings are analysed in the light of predictions made by three popular models of second language acquisition (SLA), namely Feature Model (Brown 1998, 2000), Perceptual Assimilation Model (Best 1995) and Speech Learning Model (Flege 1995), to see which of these models can better explain the acquisition of L2 consonants of English by those learners whose L1 is Pashto.

The findings show that the vowel context and L1 features play a significant role in the acquisition of L2 sounds. Although the findings of this study confirm a correspondence between perception and production of consonants, certain factors influence production and perception differently. The study concludes that although these models of SLA address most of the problems of the learners, none of these models comprehensively encompasses all the factors involved in L2 acquisition. Thus an integrated approach based on the predictions of all the models of SLA is recommended for thoroughly understanding the problems faced by L2 learners.

Keywords: English consonants; Pashto speakers/learner; Feature Model; Perceptual Assimilation Model; Speech learning Model

1. Introduction and Background

Pashto (also called Paxto) is one of the major languages of Pakistan. For the ease of our own understanding, we divide it into two dialects in Pakistan, the northern and the southern Pashto. Besides other dialectal differences, one important difference between the Pashto of the north and that of the south is that where the southern Pashto uses palatal fricative sound [ʒ], the northern Pashto uses velar stop [g] in certain words. All of the participants of this study were from the Khyber Pakhtunkhawa Province (formerly NWFP) where [g] is preferred. The following list of consonants of Pashto has been adopted and modified from Elfenbein (1997:733).

Table 1:
Phonemic Inventory of Pashto

	Labial	Dental	Palato-Alveolar	Retroflex	Velar	Uvular
Stops	p b	t̪ d̪		ʈ ɖ	k g	
Affricates		c ɟ	cˇ ɟˇ			
Fricative					x ɣ	h
Spirants		s z	ʃ ʒ	xˇ gˇ		
Nasal	m	n		ɳ		
Flaps		r		ɽ		
Lateral		l				
Semi-Vowel	ʋ [1]	j				

In order to have a clear understanding of any study on the linguistics of Pakistan, the knowledge of the educational and linguistic scenario of Pakistani society is necessary. Urdu is the national language of the country and is dominant in media and educational institutions. As Pakistan is a multi-linguistic country, we need a common language in the country. Thus, Urdu enjoys the status of *lingua franca* in Pakistan. As a result, every school going student learns Urdu. Consequently, any student who starts learning English in Pakistan has already learnt Urdu.

Urdu has [f] in its phonemic inventory and it also has four-way contrast based on voicing and aspiration. The Pashto speakers of Urdu are normally found to confuse [f] with [p] and the aspirated stops with the unaspirated ones. This confusion between [f] and [p] or between aspirated and unaspirated stops has become a shibboleth for the Pashtoon speakers of Urdu. The speakers

themselves are aware of this and they consciously try to acquire proper [f] and aspirated stops. It has also been observed that under the stream of hyper-correction, [p] is sometimes uttered as [f], and the unaspirated stops are uttered as aspirated by the speakers whose L1 is Pashto.

This paper has been divided into four parts. In part first, literature related to the current study has been reviewed and research questions have been raised. Part second is the detailed discussions of perception test and part three discusses in detail the findings of the production test. The essay ends with the concluding remarks and recommendations for future research.

1.1. Literature Review

Brown (1998, 2000) predicted that the level of difficulty in production of a particular sound of L2 depends on the Feature Geometry of the L1 of the learner. If a particular feature which distinguishes one sound from the other is active in the L1 of the learner, the contrast will be easier to perceive though the sounds do not exist in L1 of the learner and if the particular feature is not active in the L1 of the learner, the sound will be difficult for the adult L2 learner. The model suggested by Brown is called Feature Model (hereinafter called FM) of Second Language Acquisition.

Perceptual Assimilation Model (herein after called PAM) of Second language Acquisition by Best (1994, 1995, 2007) gives a directionality of difficulty in the perception of L2 sounds by dividing all the sounds of L2 into the following categories:

1) **Two-Category (TC) Type**: Two sounds of L2 may be considered 'gesturally similar' to two native phonemes of L1 (Best 1994:191);
2) **Single Category (SC) Type:** Two sounds of L2 may be considered equally similar/discrepant' to one native phoneme of the learner (Best 1994:191);
3) **Category Goodness (CG) Type:** Two sounds of L2 may be considered similar to a single L1 sound but one may be considered good exemplar than the other of the native phoneme.
4) **Categorized-Uncategorized (CU) Type:** One of a pair of L2 sounds may be considered similar to a native phoneme while the other may be considered a different/uncategorized sound (Best & Tyler 2007).
5) **Uncategorized-Uncategorized (UU) Type:** Each of a pair of L2 sounds is considered different from the L1 sounds and hence uncategorized (Best & Tyler 2007).
6) **Nonassimilable (NA):** The L2 sounds considered 'non-speech sounds' by the learners (Best 1994:191).

Following is the directionality of difficulty for adult L2 learners in the discrimination of those L2 sounds which may be categorized with the native phonemes:

SC > CG < = > UC > TC

It means the SC type of sounds will be the most difficult for the L2 learners and the TC type will be the easiest ones. CG and UC type sounds are easier than SC but more difficult as compared to TC type of sounds.

Flege's (1995) Speech Learning Model (hereinafter SLM) divides the sounds of L2 into 'identical' 'similar' and 'new' and forwards seven hypotheses about different situations which are reproduced in the following lines:

1. The L2 learners are sensitive to the allophonic variance of the sounds of L2.
2. A new phonetic category for a non-native sound may be established if the learners perceive a little difference between the L1 and the nearest L2 sound.
3. Age of Learning (AOL) and perceptibility of differences in L2 sounds are in inverse proportion.
4. Distance between corresponding sounds of L1 and L2 is proportional to learnability which means more distant sounds are learnt more easily and vice versa.
5. Equivalence classification blocks phonetic category formation if a non-native sound is equated with a native sound.
6. L2 learners' category is deflected away from the native speakers.
7. There is a correspondence between perception and production of the L2 sounds by the learners.

Wester et al (2007) studied the production of English [ð] and [θ] by Dutch speakers. For the Dutch learners these sounds are non-native. They hypothesized that if the learners substitute these sounds with [z s] which are phonologically closer to these sounds, it means that the learners substitute sounds on phonological basis and if they substitute these sounds with [v f] which are phonetically closer to these sounds it means the learners substitute sounds on phonetic basis and if they substitute these sounds with [t d] which are relatively more marked it means the learners substitute sounds on the basis of markedness. The findings of their experiment show that the learners substitute sounds on the basis of markedness and phonology more than on the basis of phonetic cues.

Chang, Plauche & Ohala (2001) observed that the learners misperceived [k] as [t] but not vice versa. They tried to know the nature and reason of perception of [k] as [t]. They were interested to know if this substitution is because of markedness. They found that the uni-directional perceptual confusion between [t] and [k] is triggered by acoustic cues and context (front vowel 'I' after the consonant) not by markedness of velars vis-à-vis universally unmarked coronal sounds. The only acoustic difference between [k] and [t] is that a spectral peak which is distinctive between these two sounds is neutralized due to raising of F2 on account of following high front long vowel [i:]. The result is confusion between the two. But [t] is never considered as [ki] because "listeners are unlikely to erroneously insert a non-existent cue into the speech signal for [ti]" (Ibid:81).

1.2. Research Questions

The current study is focussed on the analysis of perception and production of some sounds of English by such Pakistani learners/speakers whose mother tongue is Pashto. Perception test was focussed on [f], [v], [w], [θ], [ð], [t^h], [k^h], [p^h], [ʒ], [s], [z], [j], [ʧ] and [ʤ]. Both Pashto and English have [p], [t] and [k] but in Pashto these sounds are not used in aspirated form while in English they are used in both aspirated and unaspirated form depending on the context. Thus both aspirated and unaspirated allophones of these stops were part of the study in the perception test. Production test was focussed on [f], [v], [w], [θ], [ð], [ʒ], [ʧ] and [ʤ] only.

By having a look at the phonemic inventory of Pashto, we realize that the English sounds [f], [v], [w], [θ], [ð] , [t^h], [k^h], and [p^h] are non-existent in Pashto. However, we can hypothesize that affricates of English [ʧ] and [ʤ] will be equated with the corresponding native affricate sounds of Pashto by the participants. Thus, these sounds lie in the category of 'identical' according to SLM. These sounds may be easier to perceive and produce for the Pashtoon learners. Other English sounds like [θ], [ð], [f], [w] and [v] may be either 'new' or 'similar' for the learners. The English sound [ʒ] presents an interesting case. This sound is substituted with [g] in most of the areas from where the participants were. Thus this sound exists in another dialect of Pashto but not in the dialect spoken by the participants. SLM predicts the sensitivity of the learners about the allophonic variance in L2 but no model of SLA explicitly discusses the impact of allophonic variance of L1 on the acquisition of L2. The variation of [ʒ]/[g] sounds will provide an interesting context for such a study.

The English [θ] and [ð] lie closer to Pashto [s] and [z] respectively in terms of manner of articulation and they are closer to Pashto sounds [ʈ] and [ɖ] in terms of place of articulation. If the learners perceive these sounds as [s] and [z], these sounds already exist in both Pashto and English. In such a situation, the English sounds [θ s] and [ð z] will make either CG or UC type in terms of PAM but if they are perceived [ʈ, ɖ] on the basis of their of place of articulation, they will lie in TC type and will be relatively easier to perceive. The aspirated and unaspirated pairs of stops may also be expected to lie in the CG/UC type in the L2 phonemic inventory of the participants. It is expected that the participants will equate the English sounds [v] and [w] with the native Pashto labio-dental approximant thus making an SC type of pair in terms of PAM. The English sound [f] may also be either equated with [s] on account of manner of articulation or with [p] on account of place of articulation.

The current study aims to analyze the perception and production of English stops, affricates, dental and palatal fricatives and approximant. It will address the following questions:

1. How do the Pashtoon learners of English perceive and produce non-native sounds of English?
2. Is the directionality of difficulty faced by the Pashtoon learners of English in accordance with the predictions of PAM?
3. Does vowel context have any influence on the perception and production of L2 sounds of English by the learners whose L1 is Pashto?
4. Is there any correspondence, as predicted by SLM, between perception and production of non-native sounds of English by the speakers of Pashto?

2. Perception Test

2.1. Research Methodology for Perception Test

The study is based on two experiments, namely perception and production tests. The participants for both of the experiments were same. 7 male adult speakers of English whose mother tongue was Pashto participated in these experiments. All but one of the participants had started learning English in Government Schools in Pakistan between the ages of approximately 10-12 years. One of the participants started learning English at the age of 5 years. None of them had studied in English medium schools. All were university teachers in Pakistan. All of them had got scholarship (funding) for Ph.D. from Government of Pakistan through the universities where they had been employed as teachers. Five participants were

from the Department of Mathematics and two from Computer Science in the University of Essex. The detail of the participants is given in Appendix A.

2.1.1. Stimuli

CVC non-word stimuli were recorded in a quiet Psycholinguistic Laboratory in the Department of Language and Linguistics, University of Essex in the voice of a female native speaker of English (aged 27) using M-AUDIO MICROTRACK II Professional 2-Channel Digital Audio Recorder which was set on default setting of 44.1 KHz Sample rate and 16 Bit Depth. The recordings were transferred into computer in WAV format. Vowels before and after the C in the VCV structure were the same. Each phoneme was recorded in the context of three quantum vowels [u:], [i:] and [a:]; for example the consonant [s] was recorded like [*eesee*], [oosoo], [aasaa] etc. However, it was rather impossible for the native speaker to utter unaspirated stops i.e. [p, t, k] in such context. So, she spoke sp/t/k clusters followed by long vowels which were later on edited into CV syllables by deleting /s/ of the clusters. Audacity 1.3 Beta (Unicode) Software was used for editing. The stimuli were presented to five native speakers for verification before using for experiment and they verified that the sounds (all except the edited ones) were in quite understandable native English accent. In case of edited sounds, all of the native speakers identified these sounds as voiced ie [p t k] were identified as [b d g] because they could not expect unaspirated stops in the syllables except in coda position or in clusters. Ethical approval for this study was taken from the University of Essex and the written consent for using the data for research purpose was also taken from the participants.

2.1.2. Perception Test

The tests were conducted in the quiet rooms. Before starting the test, the nature of the tests was explained to the participants in Urdu. The stimuli saved in WAV format files were played in continuity with 4-5 seconds ISI in Audacity 1.3 Beta (Unicode) Software. However, the participants were allowed repetition as many times as they wanted. They were instructed to note the consonant on the provided answer sheet only when they had fully identified it. They were asked to use English as well as Urdu or Pashto letters to note their responses (which were soon after the test translated into English in the presence of the participants). Thus they noted their responses in Urdu and/or Pashto besides English. This technique was adopted because English does not have commonly known letters for some of its sounds.

2.2. Results of the Perception Test

The overall accuracy of the participants in the perception of the target consonants is given in Table 2 followed by the graph (Figure 1) reflecting the same data.

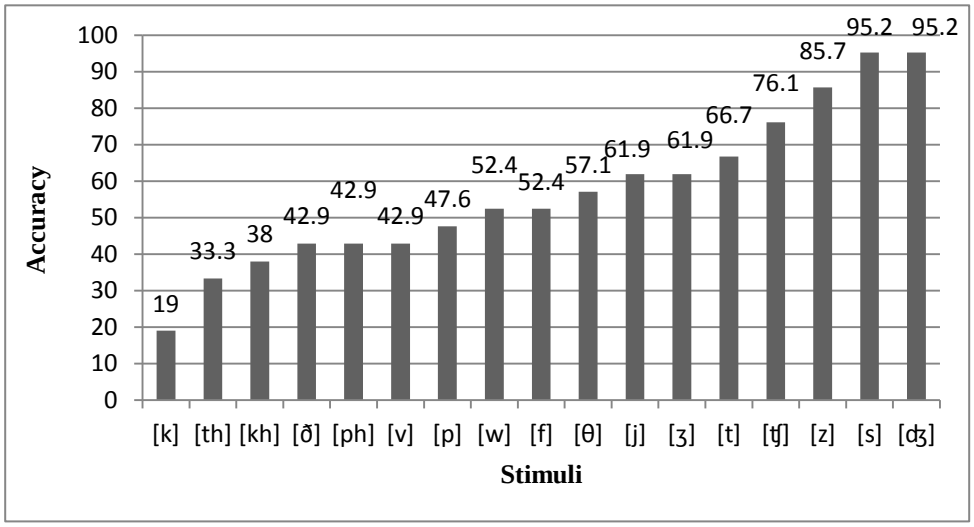
Table 2:

The Chart showing accuracy of the participants in Perception Test

Stimuli	Accuracy (in %)
[k]	19.0
[t^h]	33.3
[k^h]	38.0
[ð]	42.9
[p^h]	42.9
[v]	42.9
[p]	47.6
[w]	52.4
[θ]	57.1
[j]	61.9
[ʒ]	61.9
[t]	66.7
[ʧ]	76.1
[z]	85.7
[s]	95.2
[ʤ]	95.2

Figure 1:

Graph Showing %age Accuracy in the Perception of Consonants



Tables 3 & 4 show the detailed results of the perception test followed by Figures 6 & 7 which reflect the %age of consonant overlapping.

Table 3:

Results of Perception Test

R → S↓	ʤ	ʧ	ð	p	f	ʒ	s	tʰ	w	j	v	t	θ	kʰ	pʰ	k	z
ʤ	**20**					1											
ʧ		**21***															
ð			**9**						1		3						6
f				3	**11**		3						4				
ʒ	1					**14**				6							
s							**20**						1				
tʰ								**7**				13					
w									**11**		8						
j	1					6			1	**13**							
v			1			1			7		**9**						1
θ					1		8						**12**				
kʰ														**8**		13	
pʰ				9	1										**9**		
z			2			1	1										**17**

* five of these were perceived as aspirated sounds ie [ʧʰ]
R=Responses S=Stimuli

NB: Each sound was played 21 times (7 participants × 3 vowel contexts). Correct answers have been underlined in bold type. In some of the cases, the respondents could not identify the sound and put question marks which have not been included in the above table.

The results of the identification of the edited unaspirated stops by the participants are given separately in Table 5 below;

Table 4:

The Perception of Un-aspirated stops by the Participants

R→ S↓	p	f	t^h	w	t	k^h	p^h	k	g^h	g	$ʈ^h$	b	d	ɖ
p	**10**	1		5			2	2				1		
t			1		**11**						2		4*	
k					8	2		**4**	1	3	1			2

* One of these sounds was identified as [d^h]

Table 5:

Consonant Overlapping in Perception Test

Stimuli	C-1	%age	C-2	%age	C-3	%age
[ʧ]	[$ʧ^h$]	23.9				
[ð]	[z]	28.6	[v]	14.3		
[j]	[ʒ]	28.6				
*[f]	[θ]	19.0	[s]	14.3	[p]	14.3
[v]	[w]	33.3				
[θ]	[s]	38.0				
[ʒ]	[j]	28.6				
[w]	[v]	38.0				
[p^h]	[p]	42.9				
[t^h]	[t]	61.9				
[k^h]	[k]	61.9				

C: Confusion < 10% confusion not included

* It means, [f] was misperceived as [θ] for 19% of the time, as [s] for 14.3% and as [p] for 14.3% percent of the time. The remaining 52.4% of the time, [f] was identified correctly.

2.3. Discussion

We discuss the sounds in two groups; the sounds in which the performance of the participants is above 60% and those in which the results are below 60%. The participants have shown more than 60% accuracy in the perception test in cases of [ʧ], [ʤ], [s], [z], [ʒ] and [j] sounds. The affricates are part of the phonemic inventory of Pashto. English affricates [ʧ], [ʤ] and fricative [s z] are thus 'identical' sounds for the participants according to SLM. So these sounds must be easier for the learners whose mother language is Pashto. The result is quite according to the predictions.

The palatal fricative [ʒ] in the phonemic inventory of the Pashtoon learners may be classified among the 'identical' category because this sound exists in the phonemic inventory of the language and is quite common in southern dialect. Although this sound is substituted with [g] among the dialects spoken in the area from where the participants were selected, but the participants were quite aware of the nature and existence of that sound in their L1. Thus the results in this case are like those in the other consonants which lie in the category of identical sounds. If we study the accuracy of the participants with reference to vowel context, we realize that all the errors in identification of this sound occurred in the context of vowel [u:]. In context of [i:, a:] the accuracy is hundred percent. Thus we have strong reasons to assume that the learners are quite aware of this sound although they don't use it. They identified this sound accurately in two vowel contexts. But in the context of [u:], the labialization of the vowel influenced and changed the nature of the acoustic cues. This case is further explained in the discussion about the influence of the context on perception.

As long as the plosives [p, t, k] are concerned, theories of SLA predict accuracy of perception for these sounds by Pashto speakers because these sounds are part of the phonemic inventory of Pashto. The results of the participants in these cases seem amazing. We need to keep some factors in mind to account for these results which we discuss below.

Firstly, the results of the participants in the identification of [t] (66.7%) are better than in [p] (47.6%) and [k] (19%). The results in [t] are like other native sounds. The problem is somewhere in the other two plosives. Secondly, we should remember that we had edited the stimuli for these sounds from clusters. There is always a chance of loss of some of the acoustic cues in editing. Thirdly, there is something special with these sounds which we shall discuss one by one. As we

have already pointed out, [pʰ] and [f] don't exist in the phonemic inventory of Pashto but the speakers of this language are not only aware of the existence of these sounds in Urdu and English, but they rather consciously try to acquire these sounds. The reason for this special concentration of the Pashto speakers on these sounds is that they have to utter these sounds in Urdu language which is not only a lingua franca but also enjoys the status of national language. Thus an impression of 'educatedness' is affiliated to this language. Consequently, the speakers try to acquire these sounds properly and as mentioned before under the flow of hypercorrectivity the speakers of Pashto sometimes substitute [p] with [f] or unaspirated with aspirated sounds[2]. Thus for the [p] sound, the participants had two distracters, [pʰ] and [f], which was why their performance in this sound was relatively weaker. This will be further established when we discuss some of the examples of the substitution of [f] and [p] with each other. This trend of substitution of the native sounds with the non-native sounds under the current of hypercorrectivity is also apparent from the results of [ʧ] sound. This is a native sound for the participants, and in principle it must not be a trouble for them. But their result was 76.1% in this sound. The errors (23.9%) were all based on substitution of this sound with its aspirated counterpart. In Pashto although there is no aspiration contrast, but in Urdu there does exist the aspirated palatal sound [ʧʰ]. So, under hypercorrecitivity, the Pashtoon speaker may keep on substituting native sound with the non-native sounds so frequently that it may become part of their linguistic habits. The same they repeated in the case of the [p] sound.

Lastly, the case of [k] may be studied on acoustic grounds. The substitution of [k] with [t] on account of acoustic reasons is confirmed in literature (Chang et al 2001). In our case, although the errors of misperception of [k] as [t] are not only limited to the context of [i:] but the frequency of misperception of [k] as [t] or closer phonemes like [ʈʰ ɖ] is fairly greater (5) in this context than in the context of [u:] (2) or [a:] (4). However, only these substitutions make 64.7% of the all errors of [k]. Thus misperception of [k] as [t] or the closer phonemes of Pashto due to confusion of acoustic signals may be considered a strong reason for participants' not showing as much accuracy in case of [k] as is observed in [t].

Now we discuss the cases of the remaining non-native consonants which are [ð θ], [v w], [f] and the aspirated sounds. These sounds are confused with closer sounds of L1 as illustrated below in Table 6.

Table 6:

Comparison of Overlapping

[ð] → [v] = 14.3%	[v] → [ð] = 5%
[v] → [w] = 33.3%	[w] → [v] = 38%
[ð] → [z] = 28.6%	[z] → [ð] = 9.5%
[θ] → [s] = 38%	[s] → [θ] = 4.8%
[p^h] → [p] = 42%	[p] → [p^h] = 9.5%
[t^h] → [t] = 61.9%	[t] → [t^h] = 4.8%
[k^h] → [k] = 61.9%	[k] → [k^h] = 9.5%
[f] → [p] = 14.3%	[p] → [f] = 4.8%
[f] → [s] = 14.3%	[s] → [f] = nil
[f] → [θ] = 19%	[θ] → [f] = 4.8%

In Table 6 above, we see a similar pattern of errors in all cases except [v] & [w] in that all are confused with one or the other sound of L1 unidirectionally, which was found to be significant. [f], [θ] and [ð] are mostly confused with [s/θ], [s] and [z] respectively but [s/θ], [s] and [z] are not confused with [f], [θ] and [ð] so frequently. Similar is the case of aspirated stops. They are confused with unaspirated counterparts for a greater number of times than their unaspirated counterparts are with them. In terms of PAM, these pairs of sounds are making CG or UC type of category where one sound of L2 is either considered better exemplar of or categorized for an L1 sound than the other. In the above cases English [f], [θ], [ð], [t^h], [k^h] and [p^h] are poor exemplars of Pashto [s/θ], [s], [z], [t], [k] and [p] while the English [s], [z], [p], [t] and [k] are considered ideal exemplars of the corresponding Pashto sounds. However, the overlapping between [ð] and [v] is by no means stronger because both are new to the learners. In terms of PAM, the pair [ð, v] will make TC type of sounds because [ð] is being confused mostly with [z] and [v] with [ʊ] by the participants. On the other hand, the overlapping between [v] and [w] is almost equal. The participants assimilate and equate these two sounds [v w] with the single corresponding sound of L1 which is [ʊ] thus making SC type. Hence according to the prediction of PAM, following must be the directionality of difficulty for L2 sounds in descending order (from difficult to easier):

SC > CG <=> UC > TC

This formula predicts following sequence of difficulty in the current study;

[v, w] > ([ð, z], [θ, s], [s/θ, f], [f, p], [p^h, p], [t^h, t] , [k^h, k]) > [ð, v]

If we analyse the results of the participants in these cases, we can note that these predictions of PAM are quite true if we exclude the pairs [s/θ, f], [f, p] and those of plosives from this list. The mutual overlapping between [ð] and [v] is 19.3%

(14.3+5)[3], between [ð] and [z] 38.1% (28.6+9.5), [θ] and [s] 42.8% (38+4.8) and [v] and [w] 71.3% (38+33.3). Excluding stops and [s/θ f], [f, p] the directionality of difficulty in other phonemes is as per prediction of the Perceptual Assimilation Model (PAM) as below:

SC > CG <= > UC> TC

[v w] > [ð z], [θ s] > [ð v]

As long as the cases of [s/θ, f] and [f, p] pairs are concerned, we should realize that the difference between [s] and [θ] is not clear in the minds of the learners. Thus we can safely say that in the mind of the learners there was only one sound for two letters [θ] and [s], so sometimes they wrote [s] and sometimes [θ] on listening [f] sound in perception test. The respondents identified [f] as [s] or [θ] on the basis of manner of articulation whereas [f] was identified as [p] on the basis of similarity of place of articulation. The confusion between [f] and [s/θ] is 33.3% (19+14.3=33.3) whereas that between [f] and [p] is only 14.3%. So, we conclude that the manner rather than place of articulation plays stronger role in perception of a non-native sound.

Another important question is how to account for these results. If we take the present situation these pairs of sounds are in the CG/UC type of sounds and according to predictions of PAM the results should be better in this case than in case of SC type. Put together the total number of errors in the perception of [f] is 47.6% (19+14.3+14.3). Although this number is closer to 38.1% (for [ð/z]) and 42.8% (for [θ/s]), still it is highest of the three pairs which lie in CG/UC group. It is because [f] has two major distracters [s] and [p] in perception. In other words, it is a situation which is reverse of SC type. In SC type cases, two non-native sounds are equated with one native sound (as [v w] are equated with [ʋ]), but in the case of [f], a non-native sound is confused with two native sounds ([p] and [s]). This is a situation which has not been explicitly discussed in PAM at large. The findings of this study show that in such a situation, manner of articulation plays stronger role in the perception of the target L2 phoneme. Syed (2009) has already found stronger influence of manner rather than place of articulation on nasalization of phonemes in Saraiki which verifies that manner of articulation is plays more effective role in certain phonological processes than the place of articulation. Another important thing is that despite having two forces of attraction in L1 for the target L2 sound, still the level of difficulty in [f] is lower than the SC type of sounds (i.e [v] and [w] sounds).

Now we take up the case of the remaining phonemes. The pairs of aspirated and unaspirated stops also lie in CG/UC type but the percentage of errors in such

pairs is highest in the group. This is because the feature [spread glottis] is not functional in the Feature Geometry of Pashto while the features distinguishing [ð] from [z] and [θ] from [s] are active in the geometry of the L1 of the learners. Thus we conclude, following Brown (1998 etc) that the non-native sounds are easier to acquire if the features of the sounds are already active in the L1 of the participants.

The confusion between [s] and [θ] in all positions shows that no separate phonetic category has been established in the L2 phonemic inventory of the participants for [θ] and that it is confused with [s] by them. Similar results have been found by Guin et al (2000) in their study of the Japanese learners of English. In the L2 phonemic inventory of the Japanese, the pair of English sounds [θ] and [s] make UC category and the results of the experiments by Guin et al (2000) were against the predictions of PAM.

Such contexts make us realize that we need to develop an integrated approach for applying these models. As we noted here, the predictions of PAM became true only in cases where the features of the non-native sounds are already active in the geometry of L1 of the learners. PAM does not provide any formula for determining the directionality of difficulty in the pairs of sounds which lie within the same class as is the case of [f, p], [f, s] [ð, z], [θ, s] [p^h, p], [t^h, t], [k^h, k] pairs in this study. Thus it is the feature model which better accounts for this situation which indicates that we need to develop an integrated model of second language acquisition based on the predictions of all the major models of second language acquisition which accounts for the learners' performance in view of their complex social and linguistic backgrounds.

2.3.1. Impact of Context on Perception

Table 7 below shows the results of important sounds in the identification test in the context of vowels and the effect of vowel context on the participants' perception of them. The overall accuracy of the participants in the perception of the target consonants in the context of [u:] was 35.1% while that in the context of [a:] was 58.2%. In the case of palatal sound [ʒ], all the 7 participants perceived [ʒ] incorrect by the listeners in the context of [u:]. Six of them perceived it [j] and one perceived it affricate [ʤ]. Friedman test confirms influence of vowel context on perception of the participants (chi square = 4.742, p= .098) only if we fix the alpha[4] level at p=.1. The results of the perception test in case of the phoneme [ʒ] were analysed using Cochran's Q test which verified strong effect of vowel context on the results of the perception test (Cochran's Q= 12.286 p=.002).

Table 7:

Impact of Vowel Context on Identification

	U	A	I
[ð]	w(1) v (2)	z (3)	v(1) z (3)
[ʒ]	ʤ(1) j (6)		
[j]	ʤ(1) w(1) ʒ (4)		ʒ(1)
Total Errors	59	38	43
Accuracy in % age	35.1%	58.2%	52.7%

In the context of [u:], six out of seven participants perceived the sound incorrect. On the other hand almost 100% accuracy is seen in the identification of [ʒ] and [j] in the context of [a:] and [i:]. Thus we conclude that the confusion between the two consonants arises in the context of [u] because of labialization. The confusing influence of [u] on consonants is verified in the literature (Halle & Best 2003). Schmidt (1995) found that the Korean listeners identified the sounds of English accurately in the context of [a:] but they were confused when the sounds were followed by [u:] which she ascribes to the effect of labialization of the following vowel. Johnson and Babel (2010) have also found the effective role of vowel context on the perception of the learners. Iverson et al (2008) found the effect of vowels on identification of [v] and [w] by the subjects whose mother tongue was Sinhala. Thus it is in this case that the role of vowel context emerges as an important factor in the identification of consonants of L1. It also shows the behaviour of learners in such cases where a sound lies in the phonemic inventory of L1 but is spoken in a dialect other than that of the participants. This situation has not been addressed in detail in literature; neither any of the popular models of second language acquisition has ever accounted for this situation elaborately.

[ð] has been confused with [v] in the context of [u] but with [z] and [v] in the context of [a:] and [i:] vowels. It is because in the context of [u:] labialization becomes dominant and the listeners perceive [ð] as [v] or [w] which are symbols for the same sound for the participants because they are equating these two sounds of English with one sound of L1. But in the context of [a:] and [i:], there is no labialization to disturb the perception of the listeners so they perceive [ð] as [z] in most of the cases.

The English [f] is confused with [p] in the context of [u:] because of labiality of following vowel. In other contexts, it is confused with fricatives. All these examples verify strong influence of vowel context on perception of the participants.

3. Production Test

3.1. Methodology

In production test, the word-reading task was given to the participants. They were given a list of words written on paper carrying target consonants in the beginning and asked to read the words in carrier sentence 'I say ……again'. The utterances were recorded using M-AUDIO MICROTRACK II Professional 2-Channel Digital Audio Recorder. The audio recording device was set on default setting of 44.1 KHz Sample rate and 16 Bit Depth. The recordings were transferred into computer in WAV format soon after the tests. Using Audacity 1.3 Beta (Unicode) Software, the files were edited and the target words were elicited from the sentences. These words were presented to the judges (native speakers of English) for evaluation. The experiment was conducted in Colchester, United Kingdom and ethical approval for this experiment was taken from the University of Essex. The approval to use the data for research purpose was also taken from the participants.

3.1.1. Evaluation of the Production Test

Three native speakers of English (aged 64, 50 and 29) who were from Essex, were asked to evaluate the target consonants of English spoken by the participants on a five point scale ranging between 5-1 where 5 meant 'native like' 4 meant 'a little deflected away from the native speech' 3 meant 'very different from native speech but understandable' 2 meant 'hardly understandable' and 1 meant 'unintelligible'. A printed paper with all the words and target sounds with five cells for five different categories were given to the evaluators. The criterion for the assessment was also printed at the bottom of the paper. (See Appendix B for evaluation paper). The words were played one by one and the evaluators were allowed listening to the words as many times as they liked. They were asked to mark a consonant after they were satisfied that they had developed a considered opinion about the pronunciation. They were also asked to assess only the target consonants. Overall inter rater reliability was almost 64% (Cronbach' alpha=.642) among the judges. However, the inter rater reliability among the judgements in case of [v] was almost 88% (Cronbach's alpha= .877).

3.2. Presentation of Results of Production Test

Following is the result of production test. Table 8 and Figure 2 compare the results of production and perception tests.

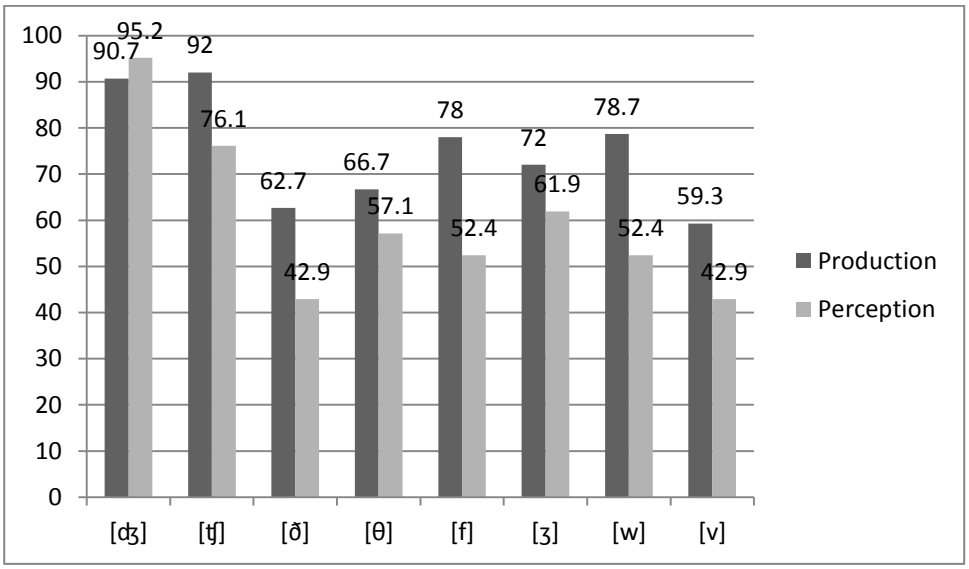
Table 8:

Perception & Production

Stimuli	Production	Perception
[ʤ]	90.7	95.2
[ʧ]	92.0	76.1
[ð]	62.7	42.9
[θ]	66.7	57.1
[f]	78.0	52.4
[ʒ]	72.0	61.9
[w]	78.7	52.4
[v]	59.3	42.9

Figure 2:

Graph showing results of Production and Perception tests



Pearson's r test showed strong correlation between perception and production test results (r=.842 p=.009 two/one tailed). A paired-sample t-test of the results of perception and production test proved the significance of the groups of data. (t-value 4.217 p=0.04 (two tailed) with Cohen's d= 0.9787818406228902 and effect size r= 0.4395740373080655).

As the perception data were categorical and the production data were in scale, Wilcoxon test was also applied to the data. The production test results were converted into two groups based on correct and incorrect perception. In one group was the result of the production test of those participants who perceived the sounds correctly and in another group were the results of those participants who has perceived the sounds incorrect. There was significant difference between the two results of the production test based on correct-incorrect perception grouping (z=10.659, N-Ties=165, p=.000. two tailed) which further confirms strong correspondence between perception and production as well as the effect of perception on the production of consonants.

3.3. Discussion

The above results show that the directionality of difficulty is from SC > CG > TC. [v w] pair is most difficult of all if we compare [ð z] [θ s] and [v w] pairs. The Pashtoon learners of English have only one sound in their language for [v] and [w]. Thus when they hear both [v w], they equate these sounds with their native sound [ʋ]. They are unable to differentiate between the two sounds. So, when they were asked to note what they had heard, they sometimes wrote 'w' and sometimes 'v' on listening [v] and [w] sounds because both were the symbols of the same sound that exists in their L2 phonemic inventory. Thus their errors in the perception of these sounds are almost equal but when they speak, for both these sounds they have one utterance and that is the same sound in their L1 which is perhaps closer to English [w] than English [v]. Thus in production test they got more marks in production of [w] than in [v]. In terms of SLM, the new category formation for [v] is blocked on account of equivalence classification. The better results in [w] are also due to equivalence classification because incidentally the two sounds in English and Pashto happen to lie very closer to each other and the positive transfer of L1 works in better pronunciation of the participants.

Correspondence between production and perception is mostly verified as following is the sequence of accuracy in both tests;

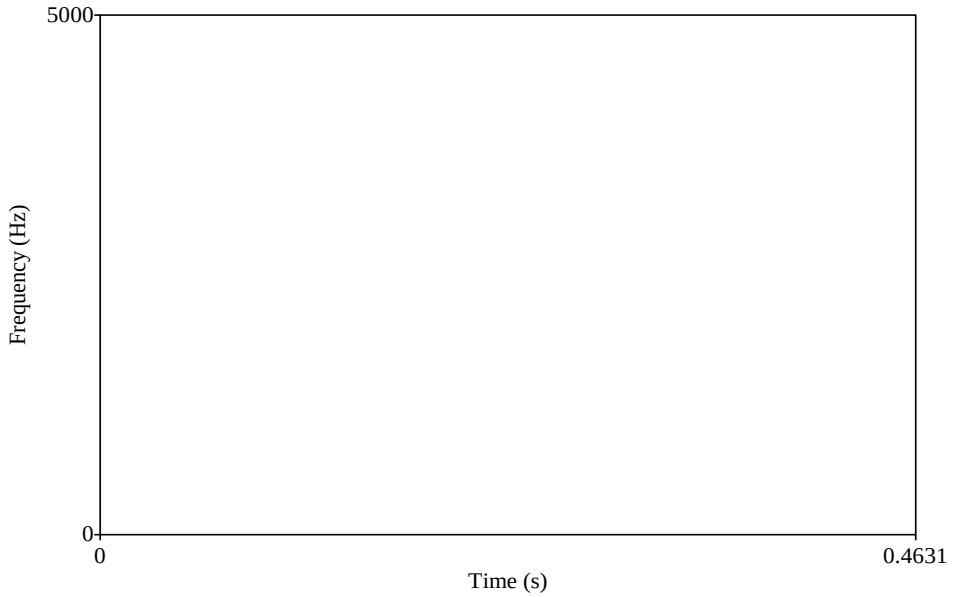
Accuracy in Perception (descending order): [ʤ ʧ] → [θ f w] → [v ð]

Accuracy in Production (descending order): [ʧ ʤ] → [f w θ] → [v ð]

The hypercorrectivity of [f] which causes substitution of [p] with [f] may be reflected from the following spectrograms taken from the words uttered by one of the participants. Figure 3 shows the spectrogram of the word 'jeep' which has been uttered as 'jeef' unconsciously by the participant and Figure 4 is the spectrogram of the word 'fall' by the same participant. The word-initial [f] has a little frication in the word 'fall' while in the word 'jeep' the word-final [p] has been substituted with [f] in such a manner that the frication of [f] is stronger here than in the word-initial [f] of the word 'fall'.

Figure 3:

Spectrogram of the word 'jeep' spoken as 'jeef'

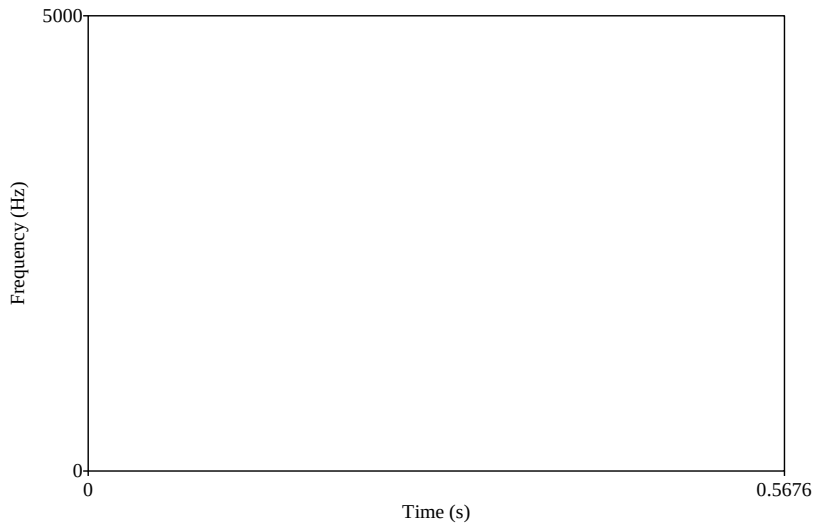


In the word 'fall' although the [f] has been uttered as fricative but the frication is not so strong as that of [f] word finally where [p] has been substituted with [f]. In this example, it is clear that the learner is consciously monitoring his speech but his focus is on word-initial consonant which is a marked position. When he speaks [f] with conscious effort, it is weak in the word 'fall' and when he utters it

unconsciously it is very strong. It means while applying models of SLA we need to account for the markedness phenomenon. The impact of L2 on L1 has also been predicted by Flege (1995:241-2) which we note in case of aspirated phonemes and [f] in Pashto. The substitution of [p] with [f] or that of unaspirated sounds with aspirated ones is example of the influence of L2 (Urdu) on L1 (Pashto).

Figure 4:

Spectrogram of the word 'fall'



We compare some examples here to analyze the real situation. Both English sounds [v] and [f] don't exist in Pashto. So these are new sounds for Pashtoon learners of English. The new phonetic category for [f] is (partially) established while that is not established for [v]. It is because of the influence of Urdu. Urdu is the lingua franca and national language. All educated Pashtoons understand and speak Urdu. Thus they are familiar with [f]. So they try to develop new category for this sound while they cannot develop a new category for aspirated sounds because of feature problem. As long as the matter of [v] is concerned, there is no social pressure on them for development of a new phonetic category for [v] because normally Pakistani languages including Urdu don't differentiate between [v] and [w] (Rahman 1990:28). Thus even the Pakistani teachers of English also cannot most of the times (or don't) differentiate between these two sounds. As a result, there is no motivation for the learners to develop a new sound. The result

is equivalence classification. Now we have three cases namely [v, w] pair, [f] and aspirated stops; the learners develop a new phonetic category for [f] because the feature fricative is active in L1 and there is a conscious effort for this on account of the influence of Urdu. In aspirated cases, the new phonetic category is not developed despite the influence of Urdu and conscious effort because the feature spread glottis is not active in the L1. The third case is [v, w] where there is neither effort, nor motivation and the result is equivalence classification although the feature is active in L1. However, it must be noted that the new category of [f] functions better in the marked environment of onset than in the coda position. It means the markedness constraint is effective even after the new phonetic category is developed.

4. Conclusion & Recommendations

At the end we summarize our findings. Our first question was how the Pashtoon learners of English perceive and produce non-native sounds of English. Our results show that the Pashtoon learners of English equate English voiced dental fricative [ð] with [v] and [z] and English voiceless dental fricative [θ] with [s]. Confusion between [v] and [w] is quite common. Aspirated stops are confused with unaspirated stops and [f] is confused with [s] and [p] in perception but with only [p] in production. Hypercorrectivity in the production of [f] leads the learners to substitute [p] with [f]. As long as the case of production is concerned, all the test sounds except [v] were produced with more than 78% of accuracy. [v] is normally confused with [w].Secondly the predictions of PAM were verified in case of directionality of difficulty but the level of difficulty within a class is still a challenge for PAM.

The third research question in the current study was if vowel context plays a role in the perception and production of the sounds of L2. It has been partially testified that the vowel context does have an influence on the acquisition of L2 sounds in case of perception. No impact of vowel context was seen on the production of sounds.

Lastly, the hypothesis by SLM about correspondence between perception and production needs separate mentioning. Overall results in the current study show strong correspondence between perception and production. The directionality of difficulty in perception and production of sounds of English is almost similar but the individual cases show a different picture. The findings of the study show that English sound [ð] is confused with [z] and [v] in perception test but it is uttered as dental stop by the participants. Similarly [θ] is perceived as [s] and [f] but produced as dental stop. [v] and [w] are mutually confused with each other

equally in the perception test but in production the performance of the learners is better in [w] than in [v]. English [f] is perceived as [s] more frequently than as [p] but in production it is confused with [p]. Palatal fricative [ʒ] is perceived hundred percent incorrect in the context of vowel [u:] and hundred percent correct in the context of other two vowels. But no difference in the performance of the participants is observed in the production test in context of vowels. Thus we conclude that the hypothesis of correspondence between perception and production of L2 sounds forwarded by SLM needs further elaboration.

Other intra-linguistic factor like location of the sounds in words (onset or coda position) extra-linguistic factors like motivation and social norms (like status of Urdu in Pakistani society in the current study) also seem to play important role in determining the level of difficulty of acquisition for the learners.

An important question is: which of the models of second language acquisition can better account for the perception and production of L2 sounds of English by those learners/users of English whose mother tongue is Pashto. We start with SLM. The sensitivity of the participants about location-specific allophonic variance of English stops is not verified in the current study. The model in its first hypothesis claims that the learners are always sensitive about the allophonic variance of the sounds of L2. But the results of the current study show that the participants cannot differentiate between aspirated and unaspirated stops. We find strong proof of equivalence classification in cases of [v] and [w]. Actually the models of second language acquisition are based on simple generalizations. For example, the SLM either gives the option of establishment of new phonetic category or equivalence classification. But there may be some stages between these two extremes because language acquisition is a gradual, complex and long process. In neither of the cases we find complete accuracy nor do we find zero percent performance in any of the phonemes. The results of both the tests in all sounds lie between 42% and 95% which means we cannot generalize anything categorically. The hypothesis of SLM that the phonemes of L2 are a little deflected away from the native categories is testified in that we don't find 100% accuracy in the production of any of the phonemes of English. As long the predictions of PAM are concerned, the principle of the directionality of difficulty is mostly verified. But within class level of difficulty is still a challenge for PAM. Similarly, FM is applicable only in cases where the features of L2 sound are involved. It gives us a wider generalization that if the features of a non-native sound are active in the geometry of L1, the sound will be easier to learn but if the features of the target L2 sound are not active in the geometry of L1, the acquisition of the sound will be difficult. But FM does not give any suggestion for

the wide variety of sounds which all lie in either one or the other class. On the basis of the findings of the current study we recommend that rather than accepting the simple (over)-generalizations of the models of SLA, the linguistic situation of a particular group of learners may be studied thoroughly from all aspects and then a particular model should be developed based on the linguistic generalizations forwarded by popular models of second language acquisition. Thus a learner specific integrated approach is needed for comprehensive understanding of the second language acquisition phenomenon.

Further research: The current study was done with a very small sample, too small to test the results statistically in most of the cases. Strong generalizations may only be made on the basis of a large representative sample. The same study may be replicated with large scale sampling taking into account various intra- and extra-linguistic factors which may influence the process of second language acquisition.

In the current study, we also came across situations not predicted by popular models of SLA. The examples of such cases are [ʒ] where one sound exists in another dialect of the L1 of the learners or there is situation which is reverse of that predicted by PAM as SC type (like [f]). In such cases we saw that the manner of articulation played an important role in case of [f] and that the vowel context was found effective in case of [ʒ]. Studies about such situations have not been discussed by popular models of SLA. Such cases need detailed longitudinal study with large number of participants.

Notes

[1] Elfenbein (1997:741) uses the symbol [w] for this consonant calling it 'Voiced labio-dental continuant' and includes it in the list of semi-vowels (ibid:742) whereas Habibullah and Robson (1996:15) call it 'voiced bi-labial semi-vowel'. We prefer to call it labio-dental approximant but unlike Elfenbein we use the symbol [ʋ] for it which is the symbol for the labio-dental approximant in IPA. The purpose is to distinguish it from English [w] and [v] for clarity in the following discussion.

[2] The researcher has been teaching English language and Linguistics for 5 years in the area where Pashto is one of the major languages and this conclusion has been drawn on the basis of personal observation.

[3] For 14.3% of the times, [ð] was considered [v] and for 5% of the times, [v] was considered [ð].

[4] Larson-Hall (2010) strongly recommends that the alpha level for the statistical tests on the L2 data must be p=0.1.

References

Best, C. (1994). 'The Emergence of Native Language Phonological in Infants: A Perceptual Assimilation Model.' In C. Goodman, & H. Nusbaum (eds.) The development of speech perception: The transition from speech sounds to spoken words, Cambridge, MA: MIT Press, 167-224.

Best, C. (1995). 'A Direct Realistic Perspective on Cross-language Speech Perception.' In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-language Research,* pp 171-206. Cambridge, MA: MIT Press.

Best, C.T. & Tyler, M. D. (2007). 'Non-native and Second Language Speech Perception: Commonalities and Complemenarities.' In *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. Amsterdam/Philadelphia, John Benjamins Publishing Company, pp 13-34.

Brown, C. (1998). 'The Role of L1 Grammar in the L2 Acquisition of Segmental Structure.' *Second Language Research* 14:136-193.

Brown, C. (2000). 'The Interrelation between Speech Perception and Phonological Acquisition from Infant to Adult.' In J. Archibald (ed.) *Second Language Acquisition and Linguistic Theory.* Malden, Mass: Blackwell, pp.

Elfenbein, J. (1997). Pashto Phonology. *Phonologies of Asia and Africa* 2 (2), pp733-759.

Flege, E.J. (1995). Second Language Speech Learning: Theory, Findings and Problems. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research,* pp 229-273, Timonium, MD: York Press.

Guion, S. G., Flege, J. E., Akahane-Ymada, 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 Acoustical Society of America* 107(5):2711-2724.

Habibullah, T. & Robson, B. (1996). A Reference Grammar of Pashto. Washton: Center for Applied Linguistics.

Halle,P. & Best C. T. (2003). 'Japanese Listeners' Perception of the French versus American English /r/ - /l/ Contrasts.' In: M.J.Sole, D. Rescasens & J.Romero (eds.) *Proceedings of 15th International Conference of Phonetic Sciences,* Barcelona, Casual production, 123-124.

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):1305-1316.

Johnson, K. & Babel, M. (2010). 'On the Perceptual Basis of Distinctive Features: Evidence from the Perception of Fricative by Dutch and English Speakers.' *Journal of Phonetics* (1):127-136.

Larson-Hall, J. (2010). A Guide to Doing Statistics in Second Language Acquisition. New York, London: Routledge

Plauche, M., Delogu, C. & Ohala, J. (1997). 'Asymmetries in Consonant Confusion.' In *Proceedings of Euro-speech '97: Fifth European Conference on Speech Communication and Technology,* 4, 2187-2190.

Rehman, T. (1990). Pakistani English: The Linguistic Description of a Non-Native Variety of English. Islamabad: National Institute of Pakistan Studies, Quaid-e-Azam University Islamabad.

Schmidt, M.A. (1995). 'Cross-language Identification of Consonants. Part 1. Korean Perception of English.' *Journal of Acoustical Society of America* 99 (5): 3201-3211.

Syed, N. A. (2009). Nasalization in Saraiki. Unpublished MA dissertation, Department of Language and Linguistics University of Essex. Copied on 30-06-2010 from website http://www.essex.ac.uk/linguistics/dissertations/2009/docs/Syed.pdf.

Wester, F., Gilbers, D. & Lowie, W. (2007). 'Substitution of Dental Fricatives in English by Dutch L2 Speakers.' *Language Sciences* (29):477-491.

APPENDIX-A

List of Participants

S. No	Age (years)	Teaching (Years)	LOR* (months)
1	35	6	42
2	32	3	15
3	35	6	24
4	27	1	19
5	38	6	27
6	37	9	27
7	35	6	42

*LOR: Length of Residence in the UK

APPENDIX-B

Evaluation Sheet

Subj: Pushto/

S.No.	Word	Target	1	2	3	4	5
1	chalk	[ʧ] ch					
2	cheese	[ʧ] ch					
3	choose	[ʧ] ch					
4	fall	[f]					
5	feel	[f]					
6	fool	[f]					
7	thaw	[θ] th					
8	thesis	[θ] th					
9	thoop*	[θ] th					
10	the's	[ð] the					
11	these	[ð] the					
12	those	[ð] the					
13	vault	[v]					
14	venus	[v]					
15	voop*	[v]					
16	wall	[w]					
17	weep	[w]					
18	wool	[w]					
19	jaw	[ʤ] J					
20	jeep	[ʤ] J					
21	juice	[ʤ] J					
22	genre	[ʒ] g					
23	zheke*	[ʒ] zh					
24	zhuke*	[ʒ] zh					

Formula	
Native-like	5
A little deflected from natives	4
Different from natives but understandable	3
Hardly understandable	2
Unintelligible	1

*Since suitable words were not available in English, these nonce words were added to the list. The participants were apprised that these were names of persons and they were also informed with examples as which sound they were expected to produce in these words.