The Sindh University Journal of Education Vol.44 No. 1, 2015, Pp.21-49

NASIR ABBAS SYED *
DR. GUL HASAN **

ROLE OF PERCEPTION IN ACQUISITION OF A FOREIGN OR SECOND LANGUAGE

Abstract

The Speech Learning Model (SLM) (Flege, 1995) predicts a correspondence between perception and production of L2 sounds (Flege, 1993). Only accurate perception leads to accurate production. The current study aims to test this prediction of the SLM. A perception and production test was conducted with a group of 11 adult learners of English. The perception test comprised of phonetic and lexical phonological identification task. The production test was based on a word reading task. In the phonetic perception task, stimuli were non-words of VCV format with English voiced stops in C position and the tense front vowel on both sides of the consonant. In the second (phonological perception) task, English words starting with voiced stops recorded in the voice of a native speaker of English were played and the participants were asked to identify those words. In the third task, the participants were asked to produce a set of English words. The productions of the participants were recorded. Voice onset time (VOT) for the voiced stops of English was measured using Praat (Boersma & Weenink, 2012). An acoustic analysis of the productions show that the participants had produced all voiced stops of English with pre-voicing. On the other hand, they perceived all labial [b] and coronal voiced stops [d] of English in the stimuli as voiceless [p] and [t], respectively. They perceived most dorsal voiced stops [g] of the stimuli as voiceless [k]. In most Pakistani languages, voiced stops are produced as pre-voiced. As a result of equivalence classification between L1 and L2 sounds, a process of negative transfer occurs in the English L2 phonemic inventory of Pakistani learners. Thus, the learners develop a representation of voiced stops of English which is based on pre-voicing. Therefore, the voiced stops produced with short-lag VOT by the native speaker of English were perceived by the participants as voiceless stops. This confirms that there is a strong correspondence between perception and production for

* Head of Department of English Language and Literature in Lasbela University of Agriculture, Water and Marine Sciences (LUAWMS), Pakistan
** PVC, LUAWMS, Lasbela University of Agriculture, Water and Marine Sciences (LUAWMS), Pakistan

these participants. The effect of vowel on adjacent stops, relationship between phonetic and phonological perception and lexical familiarity effect on perception of L2 learners was also established. The study concludes that if we have to improve our students in production of English consonants, we have to improve their perception of consonants of English Language.

Key words: *Perception, phonetics, phonology, production, VOT*

Introduction

Before last quarter of the previous century, research on second or foreign language acquisition was centered on production skill. However, by the end of the previous century, the focus of researchers changed from production to perception. Popular models of language acquisition started laying more importance to the role of perception in acquisition of a second or foreign language in adult age, e.g. see perceptual assimilation model by Catherine Best and her colleagues (Best, 1994; Best, McRoberts, & Goodell, 2001; Best, McRoberts, & Sithole, 1988; Best & Strange, 1992; Best & Tyler, 2007; P. A. Hallé, Best, & Bachrach, 2003; P.A. Hallé, Best, & Levitt, 1999), feature model by Cynthia Brown (Brown, 1997, 1998, 2000), speech learning model by James Emile Flege and his colleagues (Flege, 1992, 1993, 1995; Flege, Takagi, & Mann, 1996; Flege, Yeni-Komshian, & Liu, 1999), native language magnet theory by Patricia Kuhl and her colleagues (P. Iverson & Kuhl, 1995; P. Iverson et al., 2003; Kuhl, 1991; Kuhl, 1992a, 1992b; Kuhl, 1993, 1994; Kuhl, 2007; Kuhl et al., 2008; Kuhl & Meltzoff, 1997), cross-language assimilation overlap method by Erika S. Levy (Levy, 2009), and ontogeny phylogeny model by Roy Major (Major, 2001, 2008), etc. The current experiment aims to determine correspondence between perception and production to highlight the role of perception in acquisition of consonants of English by adult Pakistani learners.

Literature Review

This paper is mainly based on the premise that there exists a correspondence between perception and production of L2 sounds

(Flege, 1993, 1995). Thus, the main objective of this work was to confirm the correspondence between perception and production of consonants of English by Pakistani learners. However, the secondary objectives of this study were to determine the effect of lexical familiarity in L2 acquisition. Therefore, another important issue addressed in this paper was the effect of lexical familiarity on acquisition of L2 phonemes. In section 2.1, a study which provided evidence of correspondence between perception and production in acquisition of English by Chinese and Taiwanese learners was discussed. In section 2.2, another study was briefly analyzed which provided empirical support to the idea that there was a strong effect of lexical familiarity on learning of English laterals by Japanese learners.

Relationship between perception and production

The speech learning model (SLM) predicts a correspondence between perception and production of L2 sounds (Flege, 1993). It means adult L2 learners produce an L2 consonant in the same way as they perceive it. An important thing to keep in mind in this regard, is that perception precedes production (Flege, 1993, p. 1597), i.e. L2 learners acquire perception of a phoneme before its production. Therefore, in the process of learning an L2, there comes a stage when learners have acquired better perception than production. Flege (1993) studied word-final voicing contrast in English [t] and [d] with reference to vowel lengthening. He selected four groups of learners, namely less experienced Chinese adult learners (with 1.1 (s.d[1]:0.9) year of stay in the USA), less experienced Taiwanese adult learners (with 1.2 (s.d: 0.7) year of length of residence in America), more experienced Taiwanese adult learners (with 5.1 (s.d: 2.1) years of length of residence in America), and Taiwanese child L2 learners. Each group included 10 participants except for the child learners group which comprised of 9 participants. Ten native monolingual speakers of American English were also included in the study as a control group. The reason for selection of subjects from two different languages, i.e. Mandarin Chinese and Taiwanese, was that there were certain

[1]s.d. = standard deviation.

differences between these two languages which might affect learners. There are no word-final stops in Mandarin, but in Taiwanese voiceless plosives namely [p t k] are allowed to occur word-finally.

Four experiments were conducted with the participants. Each experiment had a different task: production, identification, discrimination and imitation. In the first experiment, the subjects were asked to read a carrier word of [bVt] and [bVd] structure with seven vowels of English [i], [é], [ɪ], [æ], [ɑ], [ɜ], and [u] between the two consonants. The words were produced in a carrier phrase 'I will say….' Every participant read each target word seven times. The productions were recorded. Only five middle productions were used for analysis whereas the first and last productions were discarded. In this way, 35 productions of every subject were taken for analysis. The first experiment showed that vowel duration differences in the production of a vowel before voiced and voiceless plosives by native speakers of American English averaged 137 ms but the same averaged 104 ms in productions of child learners, 63 ms in those of Taiwanese adult learners, 39 ms by inexperienced Taiwanese late learners, and 40 ms by inexperienced Mandarin late learners. These durations were taken by subtracting mean vowel duration of all productions before [d] from the mean vowel duration of all productions before [t]. The range of vowel length difference in the speech of the native speaker participants was from 84 ms to 172 ms. individually, 7 child learners, 3 experienced learners and 1 inexperienced learner were in the native-like range in production of vowels. At group level, within-group difference of vowel lengthening in the productions of the child learners was significant.

The average vowel lengthening difference in speech of the children who had arrived in the USA between the ages of 3 and 6 was 125ms (closer to the native 137 ms), but those of the children who had arrived in the USA between the ages of 9 and 13 averaged 92 ms which is far from the native duration. Post hoc comparisons confirm that the child learners and native speakers had significantly larger vowel duration differences for vowels produced before voiced and voiceless plosives. This difference was statistically non-significant in productions of the other three groups. The overall difference between vowel durations of child learners and native American speakers was non-significant ($p>.05$). Regardless of being different from native speakers, the participants produced the preceding vowels significantly longer before [d] than before [t]. Equally poor performance of both Taiwanese and Mandarin participants indicated that 'the presence versus absence of word-final stops in the L1 does not greatly influence how well adult learners will produce the English word-final /t/ vs. /d/ contrast' (Flege, 1993, p. 1593).

A second perception experiment was conducted with the same subjects. For this experiment, synthesized sounds of the words 'bat-bad' and 'beat-bead' in 17 continua were prepared as stimuli. In the stimuli, all acoustic cues of [d] were substituted with those of [t]. The stimuli were played and the participants were asked to identify which of the stimuli contained English [d] word-finally. The results of only the first three and the last three stages of the continua were analyzed. These stimuli had the longest and the shortest vowel length preceding the final consonant. The results showed that all participants had identified the stimuli with a long preceding vowel as [d] and those with a short preceding vowel as [t]. These results verify that vowel duration was an important phonetic cue for identification of the following plosives for these learners. These results confirm a correspondence between perception and production among these L2 learners. The learners had produced vowels before [t] and [d] with different vowel length in the first experiment. In second test, they perceived these sounds because of the same acoustic cues i.e. vowel length before voiced and voiceless coronal stops.

In the third experiment of the series, the same subjects were given the same 17 continua used in experiment 2 for English words 'bead' and 'beat' and the participants were asked to select the best exemplar of each word. The results showed that the native speakers and child participants based their identification of the word-final [t] and [d] on vowel duration. The experienced adult Taiwanese subjects' performance was better than that of the inexperienced Taiwanese and Mandarin learners but it was not native-like.

Experiment 4 consisted of an imitation task. The same subjects were presented the same 17 continua of the target CVC syllables (beat-bead) with the edited form of the vowel in the target words in a carrier phrase. They were asked to imitate the target words carefully. There were six repetitions in this task. The results showed that there was a great similarity between the native American English speakers group and the child learners group in this task. The remaining groups were different from these two but no significant difference was noted among them. There was a clear discontinuity in imitation response of the child and native learners but not in others which showed that the former groups had realized different vowel durations before voiced and voiceless plosives.

The findings of these experiments support the idea of a correspondence between perception and production. Those subjects who produced the vowel preceding [t d] with differing duration also perceived the consonant after the vowel on the basis of duration of the preceding vowel. This establishes a direct correspondence between perception and production of L2 phonemes by Chinese and Taiwanese learners of English. In this perspective, the aim of the current study was to confirm if there exists similar type of correspondence between perception and production of English stops by adult Pakistani learners of English.

Lexical familiarity effect on perception

Flege et al. (1996) conducted an experiment to study discrimination of English liquid consonants [l r] by Japanese speakers. One of the aims of this study was to identify the effect of familiarity on ability of learners to discriminate between two L2 phonemes. A lexical familiarity task was performed by the subjects for this purpose. The experiment consisted of two tasks. Thirty pairs of stimuli were used in the experiments. Nineteen of the stimuli consisted of words starting with [ɹ] and [l] and 4 stimuli consisted of one word and one non-word starting with each of these two English liquids i.e. [ɹ] and [l]. Seven of the stimuli consisted of words starting with [w] and [d]. The consonants [w] and [d] existed in both English and Japanese, so these sounds were used as control items. Three types of learners participated in this experiment: Native American

English speakers (NA), experienced Japanese (EJ) and inexperienced Japanese (IJ) learners of English. The stimuli were played and the participants were asked to identify them. They were asked to guess if they could not clearly understand. Ample time was given to the participants to concentrate on each stimulus and identify the sounds. It is important to point out that, the English liquids [ɹ] and [l] were very difficult to discriminate for adult Japanese learners of English because Japanese has neither of these sounds but rather has a sound which is phonetically between these two sounds.

The results of the first experiment showed that the participants identified words starting with [w] and [d] accurately. Since these were control items, it means there was no major defect in the research methodology used in this experiment. In the remaining stimuli, Japanese bilinguals misidentified [ɹ] and [l] for [w] only in 3.2% and 1.4% of the trials, respectively. The consonant [ɹ] was identified accurately by Experienced and Inexperienced Japanese participants in 92% and 76% of the trials, respectively, and [l] was identified correctly by the two groups in 77% and 63% of the trials, respectively. Among the remaining responses, bidirectional substitution of [ɹ] and [l] was seen in responses of the participants. Statistical tests confirmed significance of the difference between the two groups of participants. The results of Japanese participants were significantly better on [ɹ] than on [l].

One objective of this study was to identify the effect of lexical familiarity on discrimination of L2 learners. The results showed that lexical familiarity effect was greater on the inexperienced group than on the experienced group of participants and that it was greater for [l] than for [ɹ]. The reason for this was that [l] was used in unfamiliar or less familiar stimuli; thus, the stimuli starting with [l] consonants were found to be more difficult by the learners. The respondents mostly depended on familiarity in perception of the target consonants. Another experiment was conducted to further confirm the effect of lexical familiarity on perception of learners.

In the second experiment eight stimuli, each consisting of a pair of word-non-word sequences starting with [l] and [ɹ] were used. Four real words starting with [ɹ], four non-words starting with [l], four real words starting with [l] and four non-words starting with [ɹ] were used as stimuli in this experiment. Along with eight word-non-word stimuli, eight pairs of sounds, which were edited from these stimuli, were also included in the list. Thus, a total of 16 stimuli were used in this test, eight of which consisted of word and non-word pairs and eight consisted of edited phonetic sounds starting with [l] and [ɹ]. The hypothesis was that a lexical familiarity effect would be observed in the stimuli (words) starting with meaningful words, but not in the stimuli starting with the edited phonetic sounds. All stimuli had four randomized repetitions.

The results showed that the native American English speakers identified all stimuli (i.e. those starting with [l] and [ɹ]) in words, non-words and those in the nonce-word phonetic form) accurately. The accurate perception of native speakers of English confirmed that no change occurred in the nature of target consonants as a result of editing. The performance of the Japanese participants was also consistent with their performance in the previous experiment. The experienced learners were better than inexperienced ones, their results on [ɹ] were overall better than on [l], and they perceived lexically familiar words more accurately than phonetic productions of edited non-word stimuli. These experiments verified that lexical familiarity had carried a very significant effect on perception of L2 learners.

In the above studies, a correspondence between perception and production was established with reference to gradient phonetic differences between two types of stimuli i.e. long and short vowels before voiced and voiceless stops, respectively. The effect of lexical familiarity is also established in the study discussed above. Both these series of experiments were conducted with speakers of East Asian languages who were learners of English. The current study aimed to test the same issues with Pakistani learners of English. Moreover, a study of the relationship between phonetic and phonemic perception and effect of adjacent vowel on perception of consonants of L2 English is also one of the secondary objectives of this experiment.

Research Methodology

The Speech Learning Model (SLM) (Flege, 1995) predicts a correspondence between perception and production of L2 sounds (Flege, 1993). The current study aimed to test this prediction with adult Pakistani learners of English. A perception and production test was arranged with a group of students of English in a Pakistani University. The details of the experiment are described in the following sub-sections.

Participants

A group of 11 students of M.Phil at the Department of English language & literature Lasbela University of Agriculture, Water & Marine Sciences, Uthal Balochistan participated in this experiment. Two of the participants were female and nine were male. The average age of the participants was 31.18 (minimum= 25, maximum=38, standard deviation= 4.19) years. All of them had obtained either MA or BS(honours) in English language and/or literature after sixteen years of education. All of them had successfully completed their first semester in M. Phil English (linguistics). The participants speak multiple languages of Balochistan. The detail of their L1s is given in table 1.

Table 1: L1 of the participants

L1	No
Balochi	4
Brahvi	3
Lasi	2
Hazargi	1
Saraiki	1
Total	**11**

Stimuli

Three different tasks namely phonetic perception task, phonological perception task and production task were conducted with the students. The first author conducted this experiment. The stimuli for phonetic perception test were recorded in the voice of a female native speaker of Essex English (aged 29) who was doing PhD in socio-phonetics at the Department of Language and Linguistics University of Essex, England at the time of recording. The recording was done in a psycholinguistic laboratory of the Department of Language and Linguistics University of Essex, England.

The native speaker allowed the researcher to record and later on use her voice anonymously for research purpose. The list of non-word stimuli for the phonetic perception task consisted of VCV syllables ([ibi] [igi] [isi] [idi] [imi]) with the Vs as English tense high front vowel and the Cs as the target consonants of English. Before, using the test, the stimuli were played in front of four native speakers of English from Essex who were asked if they confirm that the consonants between the two vowels were produced correctly in native English accent. They all confirmed that all the consonants in the stimuli were produced in native English accent. The non-words given in the list were played in Audacity 1.3 Beta Unicode Software (BUS) from a Dell laptop using sound amplifiers. Each stimulus was played several times repeatedly until the participants realized and informed the author that they had heard and understood the stimuli properly. In the list of stimuli each stimulus was written three times randomly. In this way, three responses for each stimulus were obtained.

Another task arranged with the participants was a phonological perception test. A list of words comprising of the target consonants of English (i.e. [b], [d], [g]) was used as stimuli for the second task. The list of stimuli used in the phonological perception task comprised of the English words 'ball, goose, guard, do, ghee, deem, boo, dot, beak'. The stimuli were recorded in the voice of a female native speaker of English (aged 50) living in Wivenhoe, Essex.

Permission to use the stimuli for research purpose was obtained from the native speaker.

The list of stimuli had each word three times in random order. The words had three different vowels. The purpose of using stimuli with three different vowels was to determine the effect of adjacent vowels on phonological perception of the learners. Three responses for each of the target words were obtained in this task. Each stimulus was played repeatedly until the participants realized that they had heard and understood the stimulus properly.

The third test was a word-reading task. A list of the words containing 'beak, deal, geese,' was given to the participants to read in normal natural speech. Each of the target words was written three times in the list. In this way, three repetitions for each of the target sounds were obtained. The productions were recorded and used for acoustic analysis.

Data Collection

First, the willingness of the students for their participation in the experimentation was sought. The perception test was conducted in a quiet computer laboratory of the Department of English Language and Literature, Lasbela University Uthal Balochistan. Sound amplifier speakers were connected to a computer to play the file containing the stimuli. All students sat within the listening range of the speakers. According to their own statements, all participants had normal hearing. The phonetic perception task was arranged on the first day and the phonological perception test was arranged on the second day of the experiment. In perception studies, experiments are normally conducted with participants individually. However, there are examples of joint experiments with a group of participants to save time of the researcher and the participants. This was only possible if the number of participants was small. Crowhurst and Olivares (2014) and J. R. Iverson, Anniruddh, and Kengu (2008) also used the same method for data collection in their studies.

Production task was conducted with the participants in the office of the researcher. They were called one by one at a time of convenience for the researcher and the participants, and were asked to produce the list of stimuli which were recorded. The VOT of the word-initial voiced stops were obtained using Praat (Boersma & Weenink, 2012).

Reliability of Data

The reliability of the experiment was determined on the basis of consistency in repetitions. This established the stability of the test. Two different methods were employed for this depending on the nature of tasks. In the perception tests, the reliability was determined on the basis of consistency of the participants in their responses.

For example, in phonetic perception test, 100% of the participants gave the same response for all three trials of [b] and [d]. If a participant gave the same response (right or wrong perception) in all three tokens despite the fact that the tokens were presented to them in random order, it was assumed that the participant was carefully listening to the trials.

In other words, the reliability in phonetic perception of [b] and [d] was fairly 100%. In the phonetic perception of [g], 7 out of 11 participants were quite consistent in their responses. They either perceived all three repetitions accurately or all three inaccurately. Thus, the reliability in phonetic perception of [g] was considered 63.64%.

The following table shows consistency of the participants in perception of lexical phonological words which was also determined on the basis of consistency of responses. If they perceived a sound in all three tokens consistently correct or incorrect, their performance was considered a true reflection of their perception and hence, reliable.

The data in table 2 show that except for the words 'boo' and 'do', the consistency ranged between 63% and 100%. For only two stimuli namely 'ghee' and 'beak' their consistency was 63.64% but in the remaining 5 stimuli their accuracy was more than 81%. It showed that the results truly reflected perception of the participants.

The later analysis shows that poor results in 'dot' and 'boo' are because of the adjacent vowel following the target consonants and familiarity effect. Thus, in a majority of the trials, reliability of the data was up to the mark.

Table 2: 100% Consistency in phonological perception of participants

S. No	Stimulus	All correct	All Incorrect	Total	%age
1	Ball	-	9	9	81.82
2	Goose	7	3	10	90.91
3	Guard	7	3	10	90.91
4	Do	4	1	5	45.45
5	ghee	2	5	7	63.64
6	deem	-	11	11	100
7	Boo	4	-	4	36.37
8	Dot	8	2	10	90.91
9	beak	7	-	7	63.64

For determining reliability of the tokens in production test, a Cronbach's alpha reliability test was applied on the VOTs obtained from the three repetitions of each of the stimulus to determine reliability. The reliability in production test trials for [b], [d] and [g] was 57% (alpha=.567), 73% (alpha=.732), 77% (alpha=.769), respectively. It is important to point out that normally a cutoff point of 60% reliability or a reliability coefficient of 0.6 is acceptable and 70% reliability or an alpha value of 0.7 is considered excellent for these experiments (Larson-Hall, 2010; Scholfield, 1995). In these data, the reliability level for [b] is not above the cutoff point.

Results and discussion

The results of phonetic perception, phonological perception and production tasks are detailed in the following sub-sections separately.

Phonetic perception task

VCV type of stimuli with target sounds as C and high front vowel as V produced by a female native speaker of English were played and the participants were asked to identify which consonant of English they heard between the two vowels in a stimulus. The results are given in table 3.

Table 3: Accuracy in phonetic perception task

Target Sound	Mean	Percentage	Std. Deviation
[b]	0.00	00	0.00
[g]	0.91	30.33	1.14
[s]	2.64	88.00	0.67
[d]	0.00	00	0.00
[m]	2.55	85.00	0.69

There were 3 repetitions for each target sound. Therefore, a total of 33 (3 repetitions*11 participants) responses for each target consonant were obtained. The consonants [s] and [m] were included in the list of stimuli as control sounds. The mean results of the control consonants show that participants perceived these consonants (i.e. [s] and [m]) accurately in more than 85% of trials.

This confirms that the methodology used for the experiment was accurate and the results reflect actual performance of participants. A look at table 3 shows that not a single trial of [b] and [d] was perceived correctly and [g] was perceived correctly in only 30.33% of trials.

The mean values given in table 3 are out of a total of 3 because there were three trials for each stimulus and one mark was awarded for one correct perception and zero for an incorrect perception. A detail of responses of the participants is given in table 4.

Table 4: Detailed responses of participants in phonetic perception test

S. No	Response/→ Stimuli↓	[p]	[t]	[k]	[b]	[d]	[g]	Miscellaneous	Total
1	[b]	27	--	--	--	--	--	6	33
2	[d]	--	29	--	--	--	--	4	33
3	[g]	--	--	25	--	--	10	2	33

Table 4 shows that in 27 out of 33 trails, [b] was perceived as [p] and in 29 and 10 trials, [d] and [g] were perceived as [t] and [k], respectively. The remaining responses shown as miscellaneous were so varied and deviant from the actual stimuli that they might not be interpreted into a generalization.

Phonological perception test

In the phonological perception task, English words recorded in the voice of a female native speaker were played and the participants were asked to take down which English word they heard. Each word had three repetitions. The results are given in detail in table 5.

Table 5: Phonological perception task result

Stimuli	Target consonant	Mean accuracy (%age)	Std. Deviation
ball	[b]	2.55 (85.00)	0.69
boo	[b]	2.18 (72.67)	0.75
beak	[b]	3.00 (100)	0.00
dot	[d]	2.27 (75.67)	1.27
do	[d]	2.18 (72.67)	0.87
deem	[d]	0.27 (09.00)	0.65
guard	[g]	2.09 (69.67)	1.38
goose	[g]	2.00 (66.67)	1.41
ghee	[g]	1.09 (36.33)	1.22

Table 5 shows mean scores. Since there were a total three repetitions, therefore the marks were awarded out of three. If a participant perceived the target word-initial consonant accurately but s/he perceived the actual word incorrectly, the response was considered correct. For example, 'dot' perceived as '*don't*' and 'guard' perceived as 'got' or 'god' were considered correct responses because such responses showed that participants perceived the word-initial target consonants accurately.

There was no significant difference (Chi-Square=1.64, p=.44) between the mean scores of the three words starting with [g] followed by three different vowels. But there was a significant difference between the scores of the tokens starting with [b] (Chi-Square=9.25, p=.010) and [d] (Chi-Square=14.00, p=.001) followed by three different vowels. It means the effect of adjacent vowel on perception of consonants preceding the vowels was significant for labial [b] and coronal stop [d] but not for dorsal stop [g].

The question was whether the difference in the performance of the participants on phonological perception of b- and d-initial words was because of the following vowels or because of familiarity of participants with these vowels and the acoustic nature of these stops in the stimuli. For this purpose, VOTs of the stimuli were measured using Praat. The VOTs of the word-initial stops in the stimuli used for phonetic and phonological perception have also been given. The VOT values in parentheses were those of the stimuli in the phonetic perception task.

Table 6: Voiced stops perceived as voiceless stops in phonological (phonetic) perception test

S.No	Word	Voiceless[2]	VOT
1	Ball	2	17
2	Boo	2	13
3	Beak	0	11 (10)
4	Dot	8	11
5	Do	8	18
6	Deem	30	21 (33)
7	Guard	10	20
8	Goose	10	19
9	Ghee	21	48 (29)

Table 6 shows that all voiced stops in the stimuli were produced by the native speakers of English with short-lag VOT. Thus, there was no phonetic difference between the stimuli. If the learners perceived some stimuli as voiced and others as voiceless, it was not because of variance in the stimuli. It was because of some other linguistic reason which would be discussed in detail in section 5 below.

Another hypothesis was that the variance in the performance of participants may be a result of their familiarity with the words used as stimuli. The list of stimuli was shown to participants a day after the experiment and they were asked if they were already familiar

[2] In this column are given the number of times a word-initial voiced stop was perceived as voiceless (beak was perceived as 'peak' etc.). Recall there were total 33 (3 repetitions*11 participants) for each stimulus.

with those English words. Table 7 shows the number of participants who claimed that they were not familiar with a particular stimulus.

Table 7 shows that for most of the participants the words 'ghee, deem, boo' were unfamiliar English words. The other words used as stimuli were familiar for them. We would get back to these results in section 5.

Table 7: No of participants who were unfamiliar with the stimuli of phonological perception task

S. No	Word	No of participants
1	ball	0
2	goose	1
3	guard	0
4	do	0
5	ghee	9
6	deem	8
7	boo	9
8	dot	0
9	beak	0

Production test

In the production test, the participants were asked to read a list of words containing English words starting with voiced stops /b d g/. Each word had three repetitions. The productions were recorded and VOTs of the stops were taken using Praat software. Since there were three repetitions for each consonant, the averaged results are given in table 8.

Table 8: VOTs of voiced stops

Plosive	Minimum	Maximum	Mean	Std. Deviation
[b]	-96.00	.00	-50.39	41.51
[d]	-64.33	24.00	-07.33	27.85
[g]	-73.67	33.67	-30.18	33.316

Table 8 shows that overall, voiced stops of English were produced as pre-voiced by the participants with negative VOT. The mean difference between the VOTs was highly significant (F=9.794, df=2, sig=.001). Individual results showed that out of 33 trials, 9 trials of [d] and 3 trials of [g] were produced with a post-burst VOT. However, not a single participant produced all trials of even a single consonant consistently with positive VOT. This shows that in the phonemic inventory of all participants, voiced stops of English have pre-voiced representation. This would be discussed and analyzed the results in the following section.

Analysis and discussion

Before analysis, the results are summarized very briefly. The results of the phonetic perception test showed that out of 33 trials for each of the target consonants, the participants could not perceive English [d] and [b] accurately in even a single trial (see table 3). The participants perceived English [g] accurately in 10 (30.33%) trials. English [b], [d] and [g] were perceived as [p], [t], [k] in 27 (81.81%), 29(87.88%), and 25(75.76%) trials, respectively (see table 4).

The phonological perception test results showed that the participants perceived English [b] accurately in more than 66% of the trials. Only in the word 'ghee' and 'deem', the word-initial [g] and [d] were perceived accurately in only 36.33% and 9% of the trials, respectively (see table 5).

The production test results showed that overall, participants produced all voiced stops of English with pre-voicing (see table 8).

We would discuss these results one by one keeping the objectives of the study in view. The objectives of this study were to determine;

i. Effect of adjacent vowel on phonological perception,
ii. Lexical familiarity effect on phonological perception,
iii. Relationship between phonetic and phonological perception, and
iv. Correspondence between perception and production of L2 consonants.

We noticed that effect of adjacent vowels on phonological perception of labial and coronal stops of English was found significant, but the same was non-significant for velar stop [g]. We need to think why the impact of adjacent vowel was significant on labial and coronal stops but not on velar stop. From the viewpoint of articulation, there was a wider flexibility available to articulators at labial (lips) and coronal (the front of tongue) place, but that much space or flexibility was not available to the body of tongue and soft palate which articulated velar stop [g]. That was why the acoustic nature of labial and coronal consonant changes under the influence of adjacent vowel which affects the perception of listeners.

On the other hand, according to our understanding, since there was no space available to active articulator in production of [g], the acoustic nature of the consonant did not change under the influence of the adjacent vowel. Thus, the perception of listeners was not affected by change of vocalic context. The same was apparent in the current results. However, these results were not compatible with previous studies. For example, some conjugations in Czech where the choice of a vowel suffix depended on the consonants with /k, g/ always choose /u/ (Sheer, 1999, p. 209). This issue would be further investigated.

The second of the above mentioned objectives was to determine effect of lexical familiarity on phonological perception. According to self-reported statement of the participants, out of nine English words used as stimuli in the phonological perception task, three were unfamiliar to them. 9 out of 11 participants stated that the words 'ghee' and 'deem' were unfamiliar for them as English words. And for 8 participants, the word 'boo' was unfamiliar. (See table 7 for detailed results). The results (see table 5) showed that for perception of only two words, namely, 'deem' and 'ghee', the results were significantly poorer than for other words. The selection of stimuli was done very carefully with a view that if a word-initial voiced stop of a stimulus is replaced with its voiceless counterpart, it becomes another meaningful English word.

Thus, if [d] and [g] of the words 'deem' and 'ghee' (pronounced as [gi] by native speakers of English) was substituted with [t] and [k] respectively, the words become 'team' and 'key' respectively.[3] Since the words 'team' and 'key' were more familiar for the participants, they perceived the words 'deem' and 'ghee' as 'team' and 'key' respectively. Although the participants were also not familiar with the English word 'boo' but its counterpart 'poo' was also equally unfamiliar to them. Thus, they perceived this word on the basis of acoustic cues only. Overall, the results confirmed that lexical familiarity played a vital role in perception of L2 sounds. However, there were many other phonetic factors which also play a role in the perception of consonants of L2.

Another important objective of this study was to determine a relationship between phonetic and phonological perception. A look at the results of phonetic (table 3) and phonological (table 5) perception task showed a wider disharmony in phonetic and phonological perception of participants. Keeping in view the effect of adjacent vowel on perception, we only compared in table 9 the results of perception tests which had as stimuli voiced stops immediately followed by a tense front vowel.

The comparison of results given in table 9, shows that as long as voiced labial [b] was concerned, the performance of the participants was significantly different in phonetic and phonological perception. However, in perception of [d] and [g], their performance was similar.

---

[3] *It is important to note that in Pakistani English, aspiration contrast is neutralized to unaspirated stops (Mahboob & Ahmar, 2004; Rahman, 1991). Therefore, English words* $[k^hi]$ *and* $[t^him]$ *are produced and perceived as [ki] and [tim] by adult Pakistani learners.*

They performed equally poor in both cases. A non-parametric analysis showed that the difference between results of phonetic and phonological perception tests for [g] was not significant (Z=.425, sig=.671).[4]

Thus, we could safely claim that there was a direct relationship between phonetic and phonological perception of coronal and velar stops. Similar results have already been found in a study of language attrition. In (Syed, Malik, & Mangrio, 2014) we have also found that performance of participants on labial implosives was different from that of coronal and velar implosives. Since acoustic cues for labial stops were normally different from those of other stops (Ladefoged & Maddieson, 1996), the acoustic signals of phonetic and phonological material affected L2 listeners differently. However, this issue needed further research.

Table 9: Phonetic and phonological perception

Target Sound	Stimuli Phonology-phonetics	phonology (%age accuracy)	Phonetics (%age accuracy)
[b]	[bik]-[ibi]	100	00
[d]	[dim]-[idi]	9	00
[g]	[gi]-[igi]	36.33	30.33

The major objective of this study was to determine a correspondence between perception and production of L2 learners. The participants perceived all labial [b] and coronal [d] voiced stops of English as voiceless [p] and [t], respectively. They also perceived most of voiced dorsal stop [g] as voiceless [k]. The results of production test show that the participants produced voiced stops of English with negative VOT. The reason for this was that in most Pakistani languages, voiced stops are produced as pre-voiced. As a result of equivalence classification between L1 and L2 sounds, a process of negative transfer occurred.

[4] Since, all entries for [d] in phonetic perception test result are zero, a statistical test cannot be applied on these results.

Thus, Pakistani learners developed a representation of voiced stops of English which is based on pre-voicing. On the other hand, native speakers of English produce voiced stops with post-burst short-lag VOT. For reference, we reproduced the VOTs for English voiced stops in the stimuli.

Table 10: VOTs of stops used in the stimuli for phonological and phonetic perception tasks

S.No	Word	Phonology	Phonetics
1	Ball	17	
2	Boo	13	
3	Beak	11	10
4	Dot	11	
5	Do	18	
6	Deem	21	33
7	guard	20	
8	goose	19	
9	Ghee	48	29

Table 10 shows that the stimuli of the phonetic and phonological perception tasks produced by native speakers of English had short-lag VOT. Voiced stops of English produced with short-lag VOT were therefore perceived by the participants as voiceless stops. This was because, for Pakistani learners, voiced stops means stops produced with pre-voicing whereas stops produced with short-lag VOT was a voiceless stop. These results confirmed that there was a strong correspondence between perception and production of Pakistani learners of English. They perceived only pre-voiced stops as voiced stops (but those with short-lag VOT as voiceless stops) and produced the same voiced stops with pre-voicing.

However, the results of phonological perception test were not totally in accordance with the findings discussed above. They perceived the stimulus 'beak' accurately in 100% of the trials although [b] in the word 'beak' was also produced with short-lag VOT but they did not perceive it in 'peak'. They also perceived the word 'guard' accurately in 10 out of 33 (11 participants*3 repetitions) trials. However, in the remaining trails 'guard' and 'deem' were perceived as 'card' and 'team' respectively.

This demonstrates that although there was a correspondence between perception and production there were also some other factors which influence phonetic and phonological perception. These factors were robust acoustic cues which were produced with the production of consonants. These acoustic signals also had a very important role to play in perception of consonants.

The study clearly demonstrated that for accurate production, we have to train our students on perceiving these consonants accurately. In other words, more listening was required for acquiring a better speaking skill. Besides this, we have to consider other linguistic and non-linguistic factors which directly and indirectly affected acquisition of L2 consonants.

For that purpose, we also needed to study the phonetic nature of every consonant to understand and determine how effective specific acoustic signals might be in perception of specific consonants of English. All these factors jointly determined perception of L2 sounds. Acquisition of an accurate perception could lead to acquisition of an accurate production.

Summary

A perception and production experiment was conducted with student learners of English Language. The perception test had two tasks namely, phonetic perception and phonological perception. The production task had word-reading activity. The objectives of the study were to determine the effect of vowels on adjacent consonants in phonological perception, lexical familiarity effect on perception of L2 consonants, relationship between phonetic and phonological perception and correspondence between perception and production of L2 consonants.

Voiced stops of English [b d g] were target sounds in this study. The results show that L2 learners develop a correspondence between perception and production of L2 sounds. Effect of vowels on adjacent consonants and lexical familiarity effect were found to be significant on perception of L2 stops.

A significant relationship between phonetic and phonological perception was also established in this study. On the basis of these findings, it is suggested that a better perception skill may be developed in learners of English in Pakistan to enable them to produce English consonants accurately.

References

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

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. The Journal of the Acoustical Society of America, 109(2), 775-794.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. Journal of experimental psychology: human perception and performance, 14(3), 345-360.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. Journal of Phonetics, 20(3), 305-330.

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

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

Brown, C. A. (1997). The Acquisition of Segmental Structure: Consequences for Speech Perception and Second Language Acquisition. PhD, McGill University, Montreal, Quebec.

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

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

Crowhurst, M. J., & Olivares, A. T. (2014). Beyond the Iambic-Trochaic Law: the joint influence of duration and intensity on the perception of rhythmic speech. Phonology, 31(1), 51-94.

Flege, J. E. (1992). The intelligibility of English vowels spoken by British and Dutch talkers. In R. D. Kent (Ed.), Intelligibility in speech disorders thoery measurement and management (pp. 157-232). Amsterdam: John Benjamins.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. Journal of the Acoustical Society of America, 93(3), 1589-1608.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 233-277). New York: Timonium, MD.

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical Familiarity and English-language Experience Affect Japanese Adults' Perception of [ɹ] and [l]. Journal of Acoustical Society of America, 99(2), 1161-1173.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. Journal of memory and language, 41(1), 78-104.

Hallé, P. A., Best, C. T., & Bachrach, A. (2003). Perception of/dl/and/tl/clusters: A cross-linguistic perceptual study with French and Israeli listeners. Paper presented at the International Conference of Phonetic Sciences (ICPhS) 15, Barcelona.

Hallé, P. A., Best, C. T., & Levitt, A. (1999). Phonetic vs. phonological influences on French listeners' perception of American English approximants. Journal of Phonetics, 27(3), 281-306.

Iverson, J. R., Anniruddh, D. P., & Kengu, O. (2008). Perception of rhythmic grouping depends on auditory experience. Journal of Acoustical Society of America, 124, 2263-2271.

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. Journal of the Acoustical Society of America, 97(1), 553-562.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. Cognition, 87(1), B47-B57.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. Attention, Perception, & Psychophysics, 50(2), 93-107.

Kuhl, P. K. (1992a). Infants' perception and representation of speech: development of a new theory. Paper presented at the Proceedings of International Conference on Spken Language Processing, University of Aberta, Edmonton.

Kuhl, P. K. (1992b). Psychoacoustics and speech perception: Internal standards, perceptual anchors, and prototypes. In L. A. Werner & E. W. Rubel (Eds.), Developmental Psychoacoustics (pp. 293-332). Washington DC: American Psychological Association.

Kuhl, P. K. (1993). Early linguistic experience and phonetic perception: implications for theories of developmental speech perception. Journal of Phonetics, 21(1-2), 125-139.

Kuhl, P. K. (1994). Learning and representation in speech and language. Current Opinion in Neurobiology, 4(6), 812-822.

Kuhl, P. K. (2007). Is speech learning 'gated'by the social brain? Developmental Science, 10(1), 110-120.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). Philosophical Transactions of the Royal Society of Biological Sciences, 363, 979-1000.

Kuhl, P. K., & Meltzoff, A. N. (1997). Evolution, nativism, and learning in the development of language and speech. In M. Gopkin (Ed.), The inheritance and innateness of grammars (pp. 7-44): Oxford University Press.

Ladefoged, P., & Maddieson, I. (1996). The sounds of the world's languages. Oxford: Blackwell.

Larson-Hall, J. (2010). A guide to doing statistics in second language research using SPSS. New York: Routledge.

Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. Journal of the Acoustical Society of America, 126(5), 2670-2682.

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.

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

Scholfield, P. (1995). Quantifying language: A researcher's and teacher's guide to gathering language data and reducing it to figures. Clevendon/Philadelphia: Multilingual Matters Ltd.

Sheer, T. (1999). A Theory of Consonantal Interaction. Folia Linguistica, 32, 201-237.

Syed, N. A., Malik, R., & Mangrio, R. A. (2014). Language attrition is a mirror image process of language acquisition. Kashmir Journal of Language Research, XVII(1), 13-32.