The Journal of Teaching Language Skills (JTLS)
6 (3), Fall 2014, Ser. 76/4
ISSN: 2008-8191. pp. 147-165

Acoustic Analysis of advanced and intermediate Persian EFL Learners' Pronunciation of English Vowels

V. Sadeghi
Assistant Professor, TEFL
Imam Khomeini International University Qazvin
email: vsadeghi5603@gmail.com

F. Taghavi *
M.A., TEFL
Payam Noor University
email: fs.taghavi@yahoo.com

Abstract

This paper reports the results of an experimental study on non-native production of English vowels. Two groups of Persian EFL learners varying in language proficiency were tested on their ability to produce the nine plain vowels of American English. Vowel production accuracy was assessed by means of acoustic measurements. Ladefoged and Maddison's (1996) F1×F2 measurements for American English vowel space were used as reference values to be compared with the measurements obtained from the production of Persian EFL learners. The acoustic measurements revealed that learners were not able to control acoustic parameters of vowel quality, even for the more similar vowels in the two languages, in a native-like manner due to interference from their native vowel system. Spectral accuracy measures also did not progress toward more native-like values in the productions of the more proficient learners. Interestingly, the positions achieved for the more dissimilar vowels were neither close to their closest L1 counterparts nor close to their native categories. Interpreting this latter result within the Perceptual Assimilation Model proposed by Best (1994), it could be stated that learners established new phonetic categories for the vowel contrasts that were not used in their native phonological system.

Keywords: English and Persian vowel spaces, formant frequencies, language category, speech learning model, perceptual assimilation model

Received: 06/16/2014 Accepted: 02/17/2015
*Corresponding author

1. Introduction

One of the major goals of L2 learners is to pronounce the sounds of the target language without a foreign accent. However, not all learners succeed in this endeavor. Although accented production of L2 speech is related to many factors such as age and language environment, the primary influence on the nature of an individual's accent is the sound system of their native language (L1) (Flege, 1995). A number of studies have suggested that non-native speakers' experience with a particular phonetic or phonological category might contribute to their difficulty in producing and/or perceiving sound units in a second language. More specifically, results of such studies typically suggest that L2 learners have relatively greater difficulty producing and perceiving non-native contrasts that involve phonetic features dissimilar to those used in their native languages (Best, 1994; Flege, 1995; Flege, Schirru and Mackay, 2003). The goal of the present paper is to examine the production of American English vowels by Persian EFL learners. In particular, we attempt to address the degree to which Persian EFL learners' production of English vowels can be explained by their native language phonological system.

2. Literature Review

L2 learners are generally unable to pronounce the sounds of the target language without a foreign accent. Accented production and perception of L2 speech may be due to various factors including maturational factors, amount of L2 input, and interaction between L1and L2 sound systems (Flege et al, 2003).

A number of studies have explored whether the mechanisms that operate for L1 acquisition are also operative and affective as the starting age of exposure to the L2 increases (Johnson and Newport, 1989; Pallier, Bosch, & Sebastain-Galles, 1997). Results generally suggested that learners who are exposed to the target language at an earlier age (childhood) tend to be more successful than learners whose exposure starts later in life. For example, Flege Yeni-Komshian and Liu (1999) reported stronger degree of foreign accent in the speech of English utterances by Korean learners as their age of arrival in the United States increased.

The degree of an accent can also be attributed to the quality of the input learners have received. Learners who have learned the L2 through formal classroom instruction, and those who have acquired the target language in a naturalistic setting represent two distinct groups. Empirical research on acquisition of L2 sounds by learners with short length of residence in the target language community, or with non-naturalistic exposure, suggests some evidence of the limitations of native-likeness in L2 pronunciation. For example, Flege (1987) examined the production of French [u] and [y] by three groups of American English speakers varying in L2 experience, and found that two of the groups with French experience produced French [u] and [y] with F2 values that fell within the values measured in the native French productions, while the other group, with little French experience produced more English-like [u]s. Also, Bohn and Flege (1997) investigated the production of the English vowels [ɛ] and [æ] by two groups of German-speaking adults differing in English language experience as measured by length of residence in the United States (7.5 vs. 0.5 years). They found that the productions of the experienced Germans showed two distinct vowel categories, one for [a] and one for [ɛ]. However, the productions by the inexperienced Germans overlapped, suggesting that the speakers had merged the two categories. In another study, native-like pronunciation was investigated in two groups of Dutch learners of English in a formal setting (Bongaerts, 1999). It was found that the pronunciation of some of these learners was consistently judged by native English listeners to be native-like, or authentic. The result of this study and subsequent experiments (Birdsong, 2007) indicated that native-likeness could also be observed in non-naturalistic settings if three conditions are simultaneously met: high motivation, massive exposure to the L2, and intensive training in L2 perception and production skills.

In addition, previous research investigating acquisition of L2 speech sounds has clearly suggested a strong influence of the native phonological system on the production and perception of non-native sounds (Flege, Bohn, & Jang, 1997; Flege et al, 1999). Interestingly, it has been suggested that discrimination of non-native sounds can be predicted from the perceptual relatedness of non-native categories to native categories. According to Speech Learning Model (SLM), proposed by Flege (1995), L1 and L2 sounds exist in a common phonological space, and thus influence each other. As argued by Flege et al. (2003), the interaction between the two systems involves two mechanisms, namely, category assimilation and category dissimilation. An L2 sound assimilates to an L1 sound when it is perceived as an instance of the L1 sound, despite audible differences between the two sounds. However, category dissimilation happens when learners can auditorily differentiate an L2 sound from the closest L1 sound and from the neighboring L2 sound. Flege makes the assumption that under such a condition a new phonetic category for an L2 sound can be established.

The Perceptual Assimilation Model (PAM), proposed by Best (1994), also postulates that the difficulty of discrimination on non-native vowel contrasts depends on the similarity or dissimilarity of each individual vowel to listeners' native vowel categories. This model predicts four distinct assimilation patterns of L2 phonological categories to L1 phonological categories:

(1) "two-category assimilation". Two different L2 categories are perceived as equivalent to two L1 categories. The discrimination of these contrasts should be easy.

(2) category-goodness assimilation: The two categories of the L2 contrast are assimilated to the same native category; however, one category is perceived as being more similar than the other is. Thus, the discrimination could vary in degree of goodness depending on the degree of divergence of each L2 category from the L1 category, that is, they vary in category goodness. The model predicts that learners may establish a new phonological category for the more dissimilar sound.

3) single-category assimilation: The two sounds of the L2 contrast are assimilated to the same L1 sound with equal degrees of goodness. The PAM model assumes that the discrimination should be poor unless the contrast is lexically productive.

4) Uncategorized-uncategorized: The two members of a given L2 contrast are not perceived as instances of any particular L1 categories, but as in-between several L1 categories. The discrimination may be poor or good depending on the distance in the phonological space between the two sounds of the contrast. The more distant they are, the easier to discriminate they will be.

The experimental study reported here was designed to examine the production of American English vowels by non-native Persian speakers. The aim of the study was to test whether some of the predictions and hypotheses of the PAM and SLM models could be applied to EFL learners who were learning the target language in a non-naturalistic setting, more specifically, two groups of Persian EFL learners varying in English proficiency. As reviewed earlier in this section, the PAM claims that the discrimination and production of L2 contrasts can be predicted from the perceptual relatedness of L2 vowels to L1 vowels. In light of the PAM framework, predictions of potential difficulty in producing English vowel pairs by Persian learners of English were made on the basis of the four types of assimilation patterns proposed in the PAM model: two-category assimilation, category-goodness assimilation, single-category assimilation and uncategorized.

We also intended to investigate whether the Persian learners could produce English vowels accurately, and to what extent language proficiency would influence their accuracy in the production of these vowels. Vowel accuracy was assessed acoustically, in terms of the acoustic distance between vowels in the English vowel space produced by Persian EFL learners and that produced by native American English speakers. It was assumed that the high-proficient (advanced) learners would produce vowels that were acoustically closer to those produced by native English speakers than the vowels produced by mid-proficient (intermediate) learners.

3. Methods

3.1 Materials

The stimuli consisted of a corpus of nine English words containing tokens of each of the nine English vowel categories tested, namely, [i], [ɪ], [ɛ], [æ], [ʌ], [ʊ], [u], [ɑ], and [ɔ]. All stimuli were monosyllabic words in the context of stop consonant: all words started with the voiced stop [b], except “hot” [hɑt], and ended with the voiceless stop [t]. Each target word was embedded in a suitable sentence in pre final position (table 1). Pre final position helps avoid the confounding effects of boundary tones (rising and falling tones) on segmental structure (Sluijter and Van Heuven, 1996; Zhang, Nissen & Francis, 2008).

Table 1: Stimuli and context sentences of the production experiment

Target Word	***Context sentence***
beat ([bit])	Waves *beat* against the cliffs.
bit ([bɪt])	everyone needs a little *bit* of encouragement
bet ([bɛt])	She placed a *bet* on a horse called Black.
bat ([bæt])	He kept his baseball *bat* in the garage.
butt ([bʌt])	Paul quickly became the *butt* of everyone’s jokes
Bought ([bɔt])	Dan *bought* the car for $2,000
put ([pʊt])	He *put* the coffee on the table
boot ([but])	He should have got the *boot* many years ago.
hot ([hɑt])	The bar serves *hot* and cold food

Similarly a list of six Persian words containing the six Persian vowel categories, namely, [i], [e], [a], [u], [o], and [ɑ] were selected as control materials for the comparison of the Persian vowel space with those of English (as produced by Persian and American English speakers) to find possible cases of interference on an item-by-item basis. The Persian words, like the English stimuli, were embedded in suitable sentences in pre final position.

We also used Ladefoged and Maddison's (1996) F1×F2 measurements for American English vowel space as reference values to be compared with Persian and English vowel spaces produced by Persian EFL learners (table 2). The acoustic data from Ladefoged and Maddison's (1996) study reported here are from quite a broad sample of speakers from different parts of the United

Table 2: F1 and F2 frequency values of American English Vowels (adopted from Ladefoged and Maddison's (1996) study)

Vowels	F1 (Hz)	F2 (Hz)
[i]	280	2250
[ɪ]	400	1920
[e]	550	1770
[æ]	690	1660
[ʌ]	600	1170
[ɔ]	590	880
[ʊ]	450	1030
[u]	310	870
[ɑ]	710	1100

States, thus being more representative of General American English than other recent studies that provide data from a single regional variety.

3.2 Participants

A group of 40 Persian EFL learners participated in the production experiment. Their ages ranged from 19 to 30. The participants were either university students majoring in English translation or TEFL, or those who had graduated in English. The participants varied in their English proficiency, and none of them had any experience being immersed in a native English language environment. To select the homogenous sample groups for the research, initially 70 people took a TOEFL English language proficiency test, namely, Oxford Proficiency Test. The test was taken in a quiet room, and the examiner supervised the test administration. The subjects were given one hour to answer 60 questions, and then the answers were analyzed based on the criterion the test had provided. Thus, as determined by the Oxford Proficiency Test, those who scored between 30 to 47 were classified as mid-proficient learners, and those scoring between 48 to 60 were classified as high-proficient learners. Twenty participants were randomly selected from each group of mid-proficient and high-proficient learners as the final participants. The participants were all naïve as to the purpose of the experiment. Their participation was voluntary and did not imply any kind of compensation.

3.3 Procedure

Each participant was asked to read the English stimuli three times in isolation, displayed on a computer screen. Also, 20 participants were randomly selected to read the Persian stimuli three times in a separate session. There were no hesitation, distraction, and interruption during the recordings, and thus the stimuli were all recorded normally. The participants were instructed to speak naturally at a typical rate and loudness level. The recordings yielded 1080 tokens (40 participants ×9 utterances ×3 repetitions) for the English stimuli and 360 tokens (20 participants ×6 utterances×3 repetitions) for the Persian stimuli. It should be noted that, one production of the mid-proficient group could not be analyzed, leaving 1053 tokens. The participants were recorded individually in a quiet room using a digital audio recorder, Sound Blaster X- Fi 5.1, and a directional condenser microphone. The 1413 stimulus tokens were digitized at 22.05 kHz and low-pass filtered at 4.8 kHz. The output amplitude levels for each individual speaker were normalized to the maximum amplitude range.

3.4 Measurements

All acoustic measurements were made using Praat acoustic software (Boersma and Weenink, 2010). The acoustic parameters computed for each token of the first and second set of stimuli were the values of the first and the second formant frequencies (F1 and F2 in Hz). All measurements were made at the point in the vowel where the F1 reached its maximum value.

Formant frequencies were estimated by spectral peaks from the Linear Prediction Coding (LPC) coefficients. In addition, formant frequencies were calculated by locating the strongest harmonic of the formants in a Fast Fourier Transform (FFT) spectrum. Values for both measurements were compared, and in case of a slight difference between the two, the value based on the LPC was used. If there were a larger difference, the value calculated from the FFT spectrum was used.

4. Results

Table 3 shows the F1 and F2 frequency values of Persian vowels as produced by Persian speakers. As suggested by Strange (2007), Chen et al (2001), and Zhang et al. (2008), similarity relationships between L1 and L2 sounds can be established by acoustic or articulatory descriptions of sound inventories. Thus, first we will provide a simple comparison of General American English and Persian vowel inventories based on the acoustic data obtained from Table 2 and Table 3 formant frequency values. Figure 1 shows American English vowel space, as adapted from Ladefoged and Maddison's study (1996) and Persian vowel space, as computed from the measurements applied in this study.

As can be seen, the overall structure of the vowel space of American English is quite different from that of Persian. First, Persian has a smaller vowel inventory than English, with only three front vowels [i], [e], and [a], and three back vowels [u], [o], and [ɑ]. In fact, the English lax high front [ɪ] and back [ʊ] vowels as well as the central [ʌ] vowel are absent in Persian vowel inventory. In addition, there are differences in the location of specific vowels between the two languages. The production of Persian [i] is considerably lower (having higher F1) and farther front (having higher F2) compared to the American English [i]. Similarly, Persian [a] shows higher F1 and higher F2 frequencies compared to its closest equivalent in American English, [æ], suggesting that it occupies a lower and further front position in acoustic space. Persian [e], is also farther front than its closest English counterpart, that is, [ɛ], though it is higher in the vowel space compared to [ɛ]. In addition, the production of Persian [u] is farther front (less back in the sense of having higher F2) than American English [u]. Though the magnitude of the distance along the vertical axis is considerably smaller than those for [i] and [e] and [a]. [u] in Persian is also considerably lower than its English equivalent as it exhibits a much higher F1 frequency. Furthermore, Persian [a] and English [æ], though both considered as low front vowels, are sharply different in that [a] is considerably lower (having higher F1) than [æ]. The F2 values of the two vowels are quite close, however, with [a] being to a very small extent further front than [æ]. Unlike the low front vowels [æ] and [a], the back equivalents, that is [ɑ], in Persian and

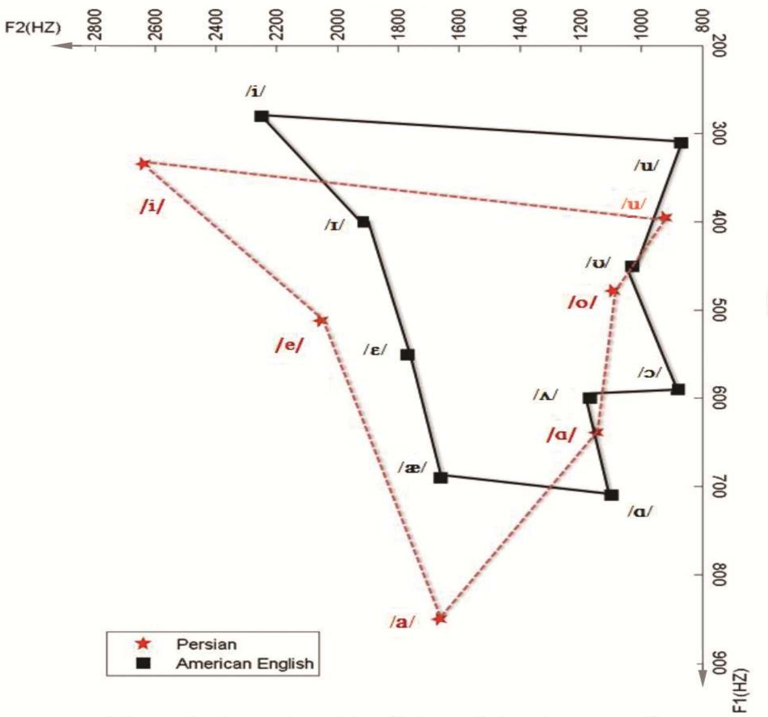

Figure1: American English and Persian vowel spaces

Table 3: F1 and F2 frequency values of Persian vowels produced by native Persian speakers

Vowels	F1 (Hz)	F2 (Hz)
[i]	334	2636
[e]	512	2053
[a]	851	1664
[ɑ]	639	1149
[o]	477	1093
[u]	395	992

English seem sufficiently close to be identified as the same vowel in the two languages. The F2 frequencies are within the same range of values, with Persian [ɑ] being slightly higher (farther front) than English [ɑ]. The F1 values seem to be relatively distant, though the magnitude of the distance is considerably smaller than that between [a] and [æ]. In addition, Persian [o] and English [ɔ] have occupied quite different positions in the acoustic space. [o] shows lower F1 and higher F2 frequencies than [ɔ]. Indeed, English [ɔ] seems to be considerably closer to Persian [ɑ] than [o], and is more likely to be identified as a vowel more similar to [ɑ] than [o] by Persian speakers.

In sum, on the basis of the comparison between the acoustic values of General American English obtained by Ladefoged and Maddison (1996) and those of Persian in the present study, it seems fair to say that the difference between English and Persian is not simply in the size of the vowel inventories, (the fact that English employs 9 simple vowels but Persian employs only 6 simple vowels). The majority of vowels in the two languages occupy quite different positions in the acoustic vowel space due to the large differences in the values of F1$\times$F2 frequencies, suggesting the possibility that the vowels concerned, though roughly regarded as equivalent, are likely to be produced with different qualities, and to be identified as different vowels by speakers of the two languages.

4.1 Learners' production of American English vowels

Table 4 shows the first and second formant frequencies of English vowels as produced by high-proficient and mid-proficient Persian EFL Learners. Nine separate one-way ANOVAs examined the effect of proficiency level on the formant frequency values of the vowels produced by the three groups of speakers. The F-values for the front vowels shown in Table 5 indicate that the simple effect of proficiency level was significant at the 0.01 alpha level for the F1 and F2 values of the [i] tokens. Fisher Post hoc tests further showed that the first and the second formants of vowels produced by both the mid-proficient and high-proficient groups were significantly higher than the F1 and F2 values of tokens produced by the American English speakers, suggesting that language proficiency did not positively influence production of the high front tense vowel [i].

Table 4: First and second formant frequencies of English vowels produced by advanced and intermediate Persian EFL learners

Vowels	High-Proficient EFL Learners		Mid-proficient EFL Learners	
	F1	F2	F1	F2
[i]	352.02	2518.15	348.48	2546.90
[ɪ]	415.91	2161.36	424.23	2306.74
[e]	636.02	1962.81	618.04	2115.98
[æ]	850.52	1734.66	807.97	1966.44
[ʌ]	681.63	1103.23	664.22	1255.96
[ɔ]	613.72	1029.08	684.41	1308.54
[ʊ]	457.45	1394.48	432.75	1272.85
[u]	376.33	1170.99	369.48	1185.10
[ɑ]	739.02	1244.73	761.22	1286.85

As for the high front lax vowel [ɪ] the F values were significant at $p<0.01$ for F2 but not for F1. The two groups of learners produced [ɪ] tokens with significantly higher F2 values than those of the native English speakers. However, the productions did not yield any significant differences between the F1 frequencies of the [ɪ] tokens produced by native and non-native speakers.

Again, the main effect of proficiency level for the mid-front vowel [ɛ] was significant at $p<0.01$ for both F1 and F2 frequencies. Both the mid-proficient and high-proficient groups differed significantly in their [ɛ] productions from the American English group as they produced the [ɛ] tokens with significantly higher F1 and F2 frequencies than those of the native English speakers according to post hoc Scheffe test.

Table 5: mean frequency values and standard deviations (in parentheses) for the vowels in the production experiment. The F-values are for separate one-way ANOVAs examining the simple effect of level of proficiency

vowel	Frequencies (Hz)	Mid-prof.	Prof.	NE	F values
/i/	F1	348.48	352.02	280	(1/62)= 79.70, ρ<0.001
	F2	2546.90	2518.15	2250	(1,62)= 86.40,

					ρ<0.001
/ɪ/	F1	424.23	415.91	400	(1/62)= 1.420, ρ=0.238
	F2	2306.74	2161.36	1920	(1,62)= 66.43, ρ<0.001
/ɛ/	F1	618.04	636.02	550	(1/62)= 46.32, ρ<0.001
	F2	2115.98	1962.81	1770	(1,62)= 54.65, ρ<0.001
/æ/	F1	807.97	850.52	690	(1/62)= 88.73, ρ<0.001
	F2	1966.44	1734.66	1660	(1,62)= 1.546, ρ=0.249
/ɑ/	F1	761.22	739.02	710	(1,62)= 28.82, ρ<0.001
	F2	1286.85	1244.73	1100	(1,62)= 31.46, ρ<0.001
/ɔ/	F1	684.41	613.72	590	(1/62)= 55.31, ρ<0.001
	F2	1308.54	1029.08	880	(1,62)= 71.63, ρ<0.001
/ʌ/	F1	664.22	681.63	600	(1/62)= 33.57, ρ<0.001
	F2	1255.96	1103.23	1170	(1,62)= 2.768, ρ=0.118
/ʊ/	F1	432.75	457.45	450	1/62)= 1.762, ρ=0.223
	F2	1272.85	1394.48	1030	(1,62)= 44.39, ρ<0.001
/u/	F1	369.48	376.33	310	(1/62)= 43.60, ρ<0.001
	F2	1185.10	1170.99	870	(1,62)= 68.74, ρ<0.001

Concerning low front vowel [a], the ANOVA revealed a significant main effect of group on F1 but not F2. Post hoc Scheffe tests indicated that the first formants for the mid-proficient and

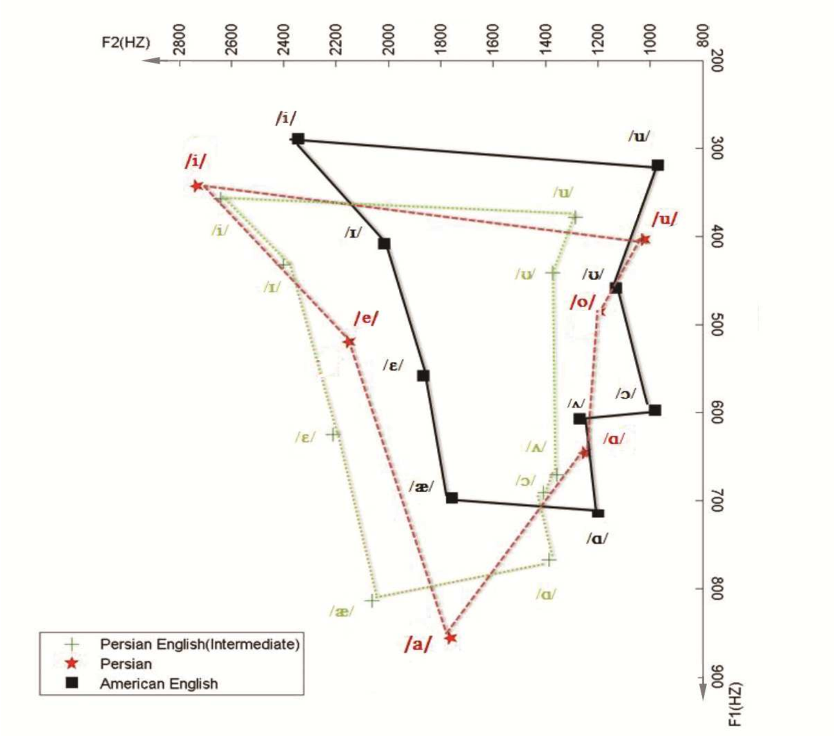

Figure 2: American English, Persian and intermediate learners vowel spaces

High-proficient groups were significantly higher than for the native English speakers. The F2 values for the two groups of learners did not differ significantly from the native English speakers group.

As regarding the back and central vowels, both groups of learners had trouble producing the high back pair [u] and [ʊ]. Significant differences were found in F1 and F2 between the [u] tokens of the native English speakers and those of the learners, regardless of their language proficiency level. Also, the two groups of learners produced [ʊ] tokens with significantly higher F2 values than those of the native English speakers. However, they did not differ significantly from the NE with respect to the first formant frequency values. For [ɔ], F values were significant for both F1 and F2 and post-hoc Scheffe tests showed that these effects were due to differences between the formant frequencies of the tokens produced by the mid-proficient and native English groups. The high-proficient learners did not differ significantly from the NE, suggesting that language proficiency positively influenced production of the back central vowel [ɔ]. This is also true for the formant values of the [ɑ] tokens, which were significantly different between NE and mid-proficient groups, but did not yield any significant differences between the native and the high-proficient productions.

For [ʌ], significant differences were found in F1 between the [ʌ] productions of the native speakers and those of the mid-proficient and high-proficient learners. However, neither group differed significantly in their [ʌ] F2 patterns from the NE group.

The mean F1 and F2 formant values for the nine vowels tested were plotted separately for each group of learners against American English and Persian (six vowels) formant values in three separate two-dimensional acoustic spaces (Figure 2 and 3). Given the large spectral, it seems that Persian learners distanced between the vowels of the NE speakers and those of the learners,

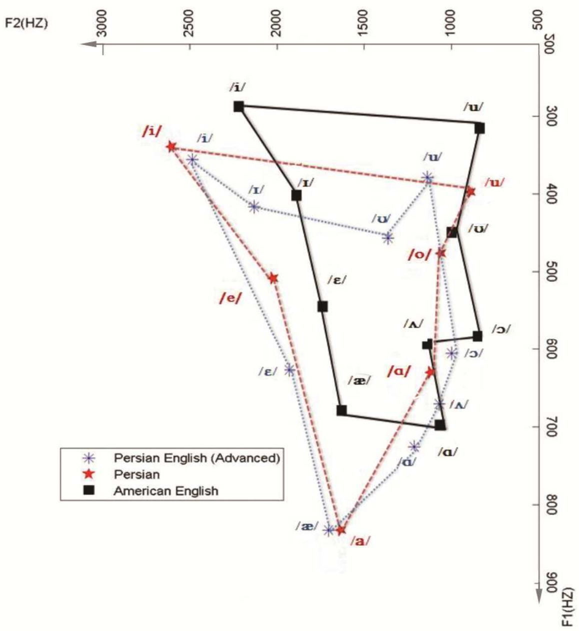

Figure 3: American English, Persian and advanced learners vowel spaces

failed to achieve the target vowel qualities required for authentic productions of American English vowels. Even in cases where the two languages employ the same vowel categories, Persian learners' productions of the target vowels do not pattern with those of the native English speakers in that they fail to produce the majority of the vowels with the expected native-like F1 and F2 values. One possible explanation for the learners' poor performance is interference from the native vowel system, or more properly, the lack of sufficiently similar vowels in the Persian system leading to particularly inaccurate production in a manner consistent with the findings of Flege et al. (1997) and Zhang at al. (2008) concerning the production of English vowels by native mandarin speakers. Persian learners' productions of American English peripheral vowels, [i], [ɪ], and [a] are acoustically more similar to native Persian vowels than their English counterparts. Just like Persian [i], Persian learners' productions of [i] are considerably lower and farther front than American English [i]. Similarly, the learners produced English [u] with much higher F1 frequency, close to the position Persian [u] occupies in the vowel space. Also, similar to Persian [a], Persian speakers' productions of American English [æ] are considerably lower (having higher F1) and further front (especially for mid-proficient learners) than [æ] produced by Native American speakers. Unlike [i], [u], and [æ], Persian learners' production of American English [ɑ] are closer to American [ɑ] than Persian [ɑ] in the vowel space. This suggests that the learners moved clearly in the expected direction of native-like [ɑ] production, though they did not manage to produce it with sufficient spectral accuracy. Persian speakers' productions of English [ɛ] are sharply different from both Persian [e], that is, its closest counterpart, and American English [ɛ]. Persian learners knew that they needed to produce a clearly different vowel quality for [ɛ], but their productions lacked sufficient accuracy to be identified as native-like [ɛ] productions.

With respect to the more central vowels [ɪ], [ʊ], [ʌ] and [ɔ], Persian learners have achieved positions that are neither acoustically close to any of the allegedly counterpart position in the Persian system, nor similar to the target points attained by their closest vowels in the American system, with which they form American English counterparts [i-ɪ], [ʊ-u], [ɔ-ɑ], [ʌ-ɑ].

This means that the learners were aware of, and attempted to make use of, formant frequency differences to achieve the central target positions required for [ɪ], [ʊ], [ʌ], and [ɔ]; however, they did not manage their productions with sufficient spectral accuracy possibly due to the lack of vowels with sufficiently similar F1 and F2 values in the Persian system.

5. Conclusion and Discussion

In general, these findings suggest that Persian EFL learners are not able to manage acoustic parameters of vowel quality in a strictly English-like manner due to interference from their native vowel system. The spectral measurements showed that the vowels produced by the two groups of learners showed a tendency to be directed towards positions occupied by Persian vowels in the acoustic space. This trend did not diminish as a function of language proficiency level, (except for [ɔ] and [ɑ]), which means that the acoustic distances between the vowels produced by EFL learners and American English speakers do not generally tend to decrease with language proficiency.

Considering the types of assimilation patterns proposed by the Perceptual Assimilation Model, it could be stated that the Persian learners followed the category goodness assimilation pattern in their productions of most American English vowel contrasts, meaning that, as hypothesized within PAM framework, they established new phonetic categories for the members of the vowel contrasts that were not perceived as similar to any L1 phonological category. Examination of the acoustic vowel spaces produced by the two groups of learners indicates that the more dissimilar members of the vowel contrasts [i,ɪ], [u,ʊ], [ɔ,ɑ], [ʌ,ɑ] (that is [ɪ], [ʊ], [ɔ], and [ʌ], which occupy central target positions untouched in L1 vowel space) have achieved new positions in the learner vowel system. The positions are neither close to their closest L1 counterparts nor close to their native categories. Thus, it may be argued that Persian adult L2 learners might establish new phonetic categories for L2 sounds if they can perceptually distinguish a given L2 sound from its closest L1 equivalent. These findings would suggest that category formation is accessible for EFL learners who learn the target language in a predominantly L1-speaking environment. Thus, this process of combining categories, known as Merger Hypothesis (Flege, 1987), may be seen as one possible source of difficulties Persian EFL learners have in pronouncing L2 vowels.

One limitation of the present study that should be addressed in subsequent in future research involves the methodological issue of how native-likeness should be investigated in L2 learning in a formal setting. The small significant differences between the two groups of learners call for an examination of production skills at the individual level, instead of the group level. This method has already been adopted by Bongaerts (1999) and Birdsong (2007). Ideally, this approach would provide a useful tool to single out the exceptional learners from the more "average" learners and thus solve the problem of within-group variability in L2 speech research.

These findings show that L1 and L2 interact to produce a hybrid system subject to modification, and offer insight into how L2 learners manage assimilation of two linguistic systems, in keeping with the Merger Hypothesis (Flege, 1987). More generally, cases of interlingual identifications found in this study has offered theoretical support for the interlanguage hypothesis assumed by Selinker (1972), and provide evidence for the reality and structural autonomy of the learner phonological systems. In fact, the observable interlingual data provided in this research may well be used for theoretical considerations concerning the psychology of the second language acquisition. As suggested by Selinker (1972), "the only observable data to which we can relate our theoretical predictions are those identified interlingually across three systems NL, TL, and IL (p. 34).

In addition, for applied considerations, as suggested by Zhang et al. (2010) the incorporation of L2 patterns of phonetic and phonological categories into course materials contributes to improving learners' pronunciation skills. Therefore, the results of this research may address experts in material development to thoughtfully and adequately incorporate native-like patterns of phonetic categories in some exercises and tasks that are intended for learning pronunciation, especially the categories which are dissimilar to L1 phonological categories. It may further be suggested that teachers explicitly make students aware of the vowel quality differences between the two languages, and employ some relevant pedagogical activities and tasks to improve learners' pronunciation of English vowels.

6. References

Best, C. T. (1994). The Emergence of Native-Language Phonological Influences in Infants: A Perceptual Assimilation Model. In: C. Goodman, & H. C. Nusbaum (Eds.). *The development of speech perception* (pp. 167-224). Cambridge, MA: The MIT Press

Birdsong, D. (2007). Native like Pronunciation among Late Learners of French as a Second Language. In: O. S. Bohn, & M. Munro (Eds.). *Language experience in second language speech learning.* In honor of James E. Flege (pp. 99-116). Amsterdam: John Benjamins.

Boersma, P., & Weenink, D. (2008). *Praat: Doing Phonetics by Computer* (Version 5.0.36) [Computer Program]. Retrieved from http:// www.praat.org/.

Bohn, O. S., & Flege, J. E. (1997). Perception and Production of a new Vowel Category by Adult Second Language Learners. In: A. James, & J. Leather (Eds.). *Second language speech : Structure and process* (pp. 53-74). Berlin and New York: Mouton de Gruyter.

Bongaerts, T. (1999). Ultimate Attainment in L2 Pronunciation: the Case of Very Advanced late L2 learners. In: D. Birdsong (Ed.). *Second language acquisition and the critical period hypothesis* (pp. 133-159). Mahwah: Erlbaum.

Chen, H., Cheung, H., Wong, O., Hills, M. (2001). *The development of phonological awareness:* Effects of spoken language experience and orthography. *Cognition*, 81(3). 227-241. Elsevier.

Flege, J. E., (1987). Effects of Equivalence Classification on the Production of Foreign Language Speech Sounds. In: J. Leather, & A. James (Eds.). *Sound patterns in second language acquisition.* Dordrecht: Foris Publications.

Flege, J. E. (1995). Second Language Speech Learning: Theory, Finding, and Problems. In: W. Strange (Ed.). *Speech perception, and linguistic experience* (pp. 233-277). Baltimore: York Press.

Flege, J. E. (2003). Methods for Assessing the Perception of Vowels in a Second Language: a Categorail Discrimination Test. In: E. Fava, & A. Mioni (Eds.). *Issues in Clinical Linguistics* 19-44. Padova: Unipress.

Flege, J. E., Bohn, O. S., & Jang, S. (1997). The production and perception of English vowels by native speakers of German, Korean, Mandarin and Spanish. *Journal of Phonetics*, 25, 437-470.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' production and perception of English vowels. *Journal of the Acoustical Society of America*, 106, 2973-2987.

Flege. J. E., Schirru, C., & MacKay, H. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communications*, 40, 467-491.

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

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The Influence of Maturational State on the Acquisition of English as a Second Language. *Cognitive Psychology*, 21, 60-99.

Ladefoged, P., & Maddison, I. (1996). *The sounds of the world's languages.* Blackwell Publishers.

Pallier, C., Bosch, L., & Sebastian-Galles, N. (1997). *A limit on behavioral plasticity in speech perception.* Cognition 64(3): 9-17. Elsevier.

Selinker, L. (1972). *Interlanguage*. International Review of Applied Linguistics in Language Teaching. 10. 209-231.

Sluijter, A. & Van Heuven, V. (1996). *Spectral balance as an acoustic correlate of linguistic stress.* J. Acoust. Soc.Am. N. 100 (4). PP. 2471-2485.

Strange, W. (2007). *Selective perception, perceptual modes, and automaticity in first- and second-Language speech processing*. Journal of the Acoustical Society of America 122.1111-1129.

Zhang, Y., & Nissen, S. L., & Francis, A.L. (2008). *Acoustic characteristics of English lexical stress produced by Mandarin speakers*. Journal of the Acoustical Society of America, 123(6): 4498-4513.

Zhang, Y., Francis, A. (2010). The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics*, 38(2), 260-271.