Perceptual assimilation of American English vowels by monolingual and bilingual learners in Iran

Afshar, N.; Heuven, V.J.J.P. van

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

Tanulmány

Naeimeh Afshar & Vincent J. van Heuven

Perceptual Assimilation of American English Vowels by Monolingual and Bilingual Learners in Iran

Abstract

In a larger project, we study the perceptual representation of the vowel system of (American) English of monolingual learners with Persian and bilingual learners with Persian and Azerbaijani as their mother tongue(s) to predict and explain problems in acquiring the pronunciation of English. In the present paper, we explore the learner's mapping of the 11 English vowels onto the 6 Persian or 9 Azerbaijani vowels within the framework of the Perceptual Assimilation Model (PAM). Twenty-two monolingual and 27 bilingual Iranian adolescent learners of English participated in this experiment. Results show that the bilinguals show different assimilation patterns than the bilinguals, even when the latter respond in Persian mode. The three extra (central) vowels in Azerbaijani may offer an advantage over the Persian system for Iranian learners of English. Neither the monolingual nor the bilingual Iranian EFL learners are sensitive to the vowel quality contrast /i~ɪ/ and /u~ʊ/ but in different ways – suggesting different learning problems.
Keywords: American English, Azerbaijani, Persian, L2 sound acquisition, Perceptual Assimilation Model, PAM test, fit index

1 Introduction

When we learn to speak a foreign language after the age of puberty, the way we pronounce the sounds of the foreign language is reminiscent of the sounds of our native language. This is what is called a foreign accent. A mild foreign accent may be charming but stronger foreign accents compromise the efficient decoding of the message and increase the risk of communication breakdown (e.g., Trofimovich & Baker, 2006; Munro & Derwing, 2008; Cutler, 2012). Most courses on English as a foreign language (EFL) contain modules which aim to improve learner's pronunciation of English, i.e., get the learner to pronounce the English sounds more like a native speaker of English would pronounce them. There is ongoing debate among experts on the question how native-like the foreigner should pronounce the English sounds (Celce-Murcia, Brinton & Goodwin, 1996; Morley, 1991; Walker, 2001), but it is generally agreed that all contrasts with a high functional load should be properly made by the foreign speaker (Jenkins, 2000, 2002; Howlader, 2010). High functional load means that there are many minimal pairs that depend on the particular contrast (Brown, 1991); if a contrast is needed to differentiate only

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

between a handful of minimal pairs, missing the contrast will not impede the speaker's intelligibility.

It is not the case that any particular sound (or contrast between two sounds) in English constitutes a learning problem *per se*. The sound system of the learner's native language determines to a large extent which English sounds may be difficult to pronounce and which ones may be easy. A systematic comparison of the sounds and sound structures of the learner's native language (L1) and the target language (L2, here English) may be used to predict relative difficulty of sounds and sound contrasts in the L2. The relevance of contrastive analysis for the acquisition of a non-native language has remained unchallenged since the original proposal was made by Lado (1957) in his Transfer Theory, although the ideas have diversified considerably in more recent decades, including the Speech Learning Model (SLM, Flege, 1987, 1995), SLM-r (Flege & Bohn, 2020), the Perceptual Assimilation Model (PAM, Best 1995, Best; McRoberts & Goodell, 2001), PAM-L2 (Best & Tyler, 2007), the Second Language Linguistic Perception model (L2LP, Escudero, 2005), and the Markedness Differential Hypothesis (Eckman, 1977, 1985).

A newborn infant is sensitive to a large number of sound categories. During the first few months of its life, the child learns to disregard the differences between many adjacent sound categories so that perceptual boundaries separating the original categories are unlearned and only the boundaries that matter in the child's ambient language remain. The most central instances of the categories, the prototypes, function as perceptual magnets: physical differences between sounds near the category prototype are perceptually underestimated and are therefore difficult – if not impossible – to detect once the perceptual categories have been set up (e.g., Kuhl & Iverson, 1995). Differences between sounds are easy to hear if the sounds are on opposite sides of a category boundary (e.g., Van Heuven & Kirsner, 2004). When we as adults hear a foreign sound, it will be drawn to the nearest available magnet in our sound system and we will not be aware of the discrepancy between the foreign sound and our native prototype that is close to it (for a summary of this position see Frieda et al., 1999; Kuhl et al., 2008).

Given this state of affairs, it is important that we establish how the sounds of a foreign language we aim to learn, are mapped onto the prototypes in our native language. The Perceptual Assimilation Model was designed to provide a typology of assimilation patterns that may occur when a listener with native language L1 is confronted with the sounds of an unknown language L2. The PA model recognizes three types of non-native sounds:

- C or *Categorized*. The foreign sound is accepted as an instance (whether good or poor) of one (and only one) of the sound categories in the native system.
- U or *Uncategorized*. These sounds fall in between two or more categories of the listener's L1. They are poor tokens of multiple adjacent categories in the L1.
- N or *Non-assimilable*. Sounds are non-assimilable when they are unlike any category in the listener's L1, i.e., are outside the phonological space of the learner's L1 (e.g., African click sounds are so unlike any consonant type in English that the English listener thinks the speaker is clapping his hands or flicking his fingers while talking, Best, McRoberts & Sithole, 1988).

This basic tripartite division of sounds is then used to describe a number of assimilation scenarios that may apply when a pair of foreign sounds has to be discriminated by a native listener of some other language. Here we will summarize four scenarios, the ones most often found:

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

- SC: *Single Category scenario*. Two different foreign sounds (i.e., contrastive phonemes in L2) are classified as equally good instances of one single category in the listener's L1. The prediction is that such a contrast in the L2 will escape the learner's attention and will constitute a persistent learning problem.
- TC: *Two Categories scenario*. Two contrastive sounds {x, y} in the L2 are assimilated in a one-to-one fashion to two contrastive sounds {a, b} in the learner's L1. The prediction is that the learner will easily discriminate between the two foreign sounds, and that the contrast will not present a learning problem.
- CG: *Category Goodness scenario*. Two different (contrastive) sounds in the L2 both map onto a single category in the learner's L1 but one matches the L1 category clearly better than the other. The listener will quickly notice the difference in goodness between the two foreign sounds, which will guide him to set up a split in his native category to accommodate the contrast (learn a new category boundary).
- UC: *Uncategorized/Categorized* scenario. One of two contrastive sounds in the L2 is assimilated to a category in the L1 while the other sound remains Uncategorized. This is like the CG scenario but the category of the less typical member is now undecided. Discrimination will be reasonable to good but the formation of the new category will be more difficult because it contains parts of multiple adjacent categories in the L1.

The primary aim of the present study is to determine how the monophthongal vowels of American English are perceptually assimilated by EFL learners in Iran. More specifically, we study two groups of Iranian learners of EFL. One group is monolingual (at the onset of EFL learning) and speaks Persian as the first and only native language. A second group was tested in the North-West of Iran near the border with Azerbaijan. At the onset of EFL learning, these learners were early bilinguals with two native languages, i.e., Azerbaijani and Persian. These learners have typically acquired Azerbaijani as their first language at home, and then learned Persian from the age of four onwards at school, where Persian is the language of instruction. When these early bilinguals participated in the present perceptual assimilation study, they were around 16 years of age (mean age = 16.9 years, against 16.5 for the monolinguals), and estimated their oral skills about equal (9 or better on a scale from 0 to 10) in both their native languages. Although quite a number of studies have been done on the perceptual assimilation of English sounds, including vowels, to the native languages of groups of monolingual learners of EFL, the perceptual assimilation by early bilingual and multilingual learners is understudied. The secondary aim of the present study is to determine whether the monolingual Persian EFL learners have different assimilation patterns than the bilingual EFL learners when the latter are instructed to map the foreign vowels onto the vowels of Persian. A related question is whether the English vowels assimilate in the same or in a different way to the vowels that are shared between Persian and Azerbaijani. Comparing the results for the bilinguals with those obtained for their monolingual Persian peers may tell us if the extra vowels in the Azerbaijani set affect the task performance in the PAM test, whether positively or negatively.

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

2 Characterization of the vowel systems involved

The monophthongal vowel system of Persian distinguishes three degrees of height (high, mid, low) and two degrees of backness (front, back). Lip rounding is unmarked, i.e., typologically normal, such that front vowels are pronounced with spread lips and back vowels with rounded lips. Persian has no contrast based on vowel duration (short, long) or tenseness (lax, tense). The approximate positions of the six vowels in the IPA vowel chart is shown in Figure 1A.

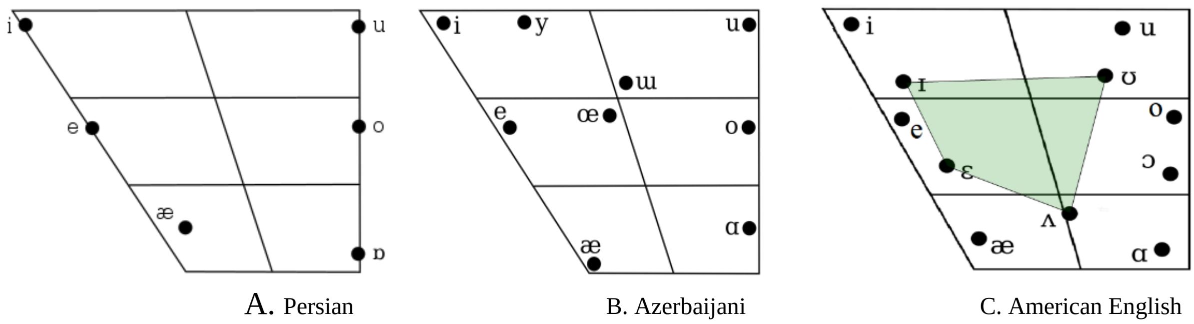

Figure 1. IPA vowel diagrams for the vowel inventories of Modern Persian (A, Majidi & Ternes, 1999), Azerbaijani (B, Ghaffarvand Mokari & Werner, 2016: 509) and American English (C, modified from Manell, Cox & Harrington, 2009; Ladefoged & Johnson, 2011: 197). The shaded quadrilateral connects the four phonetically lax/short vowels.

The vowel system of Azerbaijani is almost the same as that of Persian as far as the peripheral vowels (also called edge vowels) are concerned but it is augmented with three vowels in the central region of the vowel space, yielding a total of nine, as shown in Figure 1B. The coupling of backness and lip rounding is more complex in Azerbaijani in that the three central vowels have atypical lip rounding. Phonologically, Azerbaijani /y/ and /œ/ are front vowels (as they are in Turkish) but with (marked) lip rounding. The phonologically high /ɯ/ is a back vowel with marked spread lips. Like Persian and Turkish (closely related to Azerbaijani with a fair degree of mutual intelligibility), Azerbaijani has no length or tenseness contrast in the vowels.

The pure (monophthongal) vowel system of American English is more complex than that of either Persian or Azerbaijani. Although considerable regional variation exists, most varieties distinguish eleven vowels that are normally analyzed as monophthongs, as illustrated in Figure 1C. This system has five unrounded front vowels and five rounded back vowels, with four degrees of height: high, high-mid, low-mid and low. The monophthongs can be split into a group of seven tense vowels, and a smaller group of four lax vowels, which have shorter durations, a rather more centralized vowel quality, and no diphthongization. The tense vs. lax properties distinguish between the members of the high-mid vowel pairs /e, ɪ/ and /o, ʊ/. There is one central monophthong: mid-low /ʌ/. The (mid-) low back vowels are best analyzed as long and tense vowels as in *law* /lɔ/ and *father* /fɑðɚ/. The low front vowel /æ/ is phonetically tense (and long) in American English (e.g., Strange et al., 2004) even though it should be classified as a lax vowel on distributional grounds, since it occurs in closed syllables only and can be followed by a tautosyllabic velar nasal /ŋ/. The high-mid tense vowels /e/ and /o/ are semi-diphthongs in most varieties of English, including American English. We decided to group them with the monophthongs because the slight diphthongization is not essential for their identification, and when pronounced as monophthongs (as they are in some varieties, e.g., Scots English) they remain distinct from each other and from all other vowels – which is not the case

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

for the full diphthongs /ai, ɑu, ɔi/. Here we follow the analysis adopted by Yavaş (2011: 77–79). Also, in line with his analysis, we exclude all vowels that only occur as positional allophones before coda /r/, such as [ɚ], which is listed among the monophthongs by Ladefoged and Johnson (2011).

The auditory analyses of the vowel systems of Azerbaijani and of Persian in Figure 1A-B, have been complemented with acoustic measurements of vowel formants and duration in order to map out the phonetic details of the respective vowel spaces. Ansarin (2004) measured F1 and F2, but not the duration, in /hVd/ words spoken by 12 Persian women. Ghaffarvand Mokari, Werner and Talebi (2017) measured F1, F2, F3 and duration of the six monophthongs of Tehrani Persian (28 male, 25 female speakers) in /bVd/, /dVd/ and /hVd/ monosyllables. Aronov et al. (2017) measured vowel formants and duration in both free conversation and read-aloud word lists for two Tehran speakers of Persian. A subsequent analysis of informal Persian speech by Jones (2019) reveals that there that there is no significant length distinction between any pair of vowels (mean vowel durations between 56 and 78 ms, for /e/ and /u/ respectively), that there is regional variation in the vowel space, and that the low back vowel may be better characterized as a diphthong [ɔɒ] and is higher than often assumed. The latter finding is consistent with the earlier measurements by Ansarin (2004) and Aronow et al. (2017). The general configuration of the vowels in the acoustic space is in agreement with Figure 1A.

F1 and F2 (but not the duration) of the nine vowels of Azerbaijani spoken by 30 male and 30 female speakers were measured in /ʃVr/ monosyllables by Peivasti (2012). No breakdown by gender was presented. Ghaffarvand Mokari and Werner (2016) measured F1, F2, F3 and duration of the nine vowels of Azerbaijani produced in /bVd/ words by 20 male and 23 female monolingual speakers. The acoustic vowel plots correspond well with the traditional vowel diagram in Figure 1B. The perceptual assimilation of the vowels of Standard Southern British English (SSBE) by Azerbaijani monolinguals was subsequently studied, and predictions of perceptual and acoustic confusion of SSBE vowels by 20 male and 20 female Azerbaijani EFL learners were tested (Ghaffarvand Mokari & Werner, 2017).

Pillai and Delavari (2012) reported F1, F2 and duration in eight British English pure vowels spoken in monosyllables in a fixed carrier by 13 Persian EFL speakers (7 male, 6 female). Lack of spectral and temporal contrast was observed for the tense-lax pairs /iː~ɪ/ and /uː~ʊ/. The lax pair /ʌ~ɒ/ was conflated but the quality difference between /e~æ/ was upheld.

A contrastive acoustic study of nine monophthongs of American English (excluding the semi-diphthongs /e, o/) and the six Persian monophthongs was reported by Sadeghi and Bigdeli (2018). Tokens were sampled from existing databases of read-out continuous speech, equally divided over (an unspecified number of) male and female speakers. Formants (but not durations) of individual vowel tokens were visualized in scatterplots but no centroids or dispersion measures were provided. Subsequently, Bigdeli and Sadeghi (2020) asked 15 listeners to identify each of nine English monophthongs (three different tokens per vowel) as a vowel of Persian, with two or three response alternatives. The assimilation pattern (English > Persian) was as follows: /iː/ > /i/, /ɪ, ɛ/ > /e/, /æ/ > /æ/, /uː, ʊ/ > /u/, and /ʌ, ɔ, ɑː/ > /ɑ/. No typicality judgments were given.

Mirahadi et al. (2018) measured F1, F2 and F3 (but not duration) in the six vowels of Persian as spoken by 25 male and 25 female Azerbaijani-Persian bilingual adults. No comparison was made between bilinguals and monolingual Persian speakers.

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

There are no contrastive studies yet of the perceptual assimilation of the vowels of American English by monolingual Persian vs. bilingual Azerbaijani-Persian listeners. The present paper aims to fill this lacuna.

3 Methods

3.1 Materials

We selected two male speakers from a larger group of 20 native speakers of American English from the recordings collected by Wang and van Heuven (2006, 2014). These were the only two speakers in the set who observed a proper contrast between the (mid-)low back vowels /ɑ/ and /ɔ/. Speakers had been recorded on digital audio tape (DAT) in a sound-insulated recording booth through a Sennheiser MKH-416 microphone. Materials were later downsampled (16 KHz, 16 bits). For each speaker the following set of 11 monosyllabic words or phrases was excerpted from the fixed carrier *Now say ... again*, using the digital waveform editor in the Praat (version 6.1.05) speech processing software (Boersma & Weenink, 2019; Boersma & Van Heuven, 2001): *heed* /hid/, *hid* /hɪd/, *hayed* /hed/, *head* /hɛd/, *had* /hæd/, *hud* /hʌd/, *hod* /hɑd/, *hawed* /hɔd/, *hoed* /hod/, *hood* /hʊd/, and *who'd* /hud/, following the established practice in, e.g., Peterson & Barney (1952) and Hillenbrand et al. (1995). Acoustic details of the 22 tokens used in the PAM test can be found in the Appendix.

3.2 Participants

Two groups of listeners participated in the experiment. The first group comprised 22 native speakers (11 male, 11 female) of Modern Persian. They were secondary school pupils in Tehran with a mean exposure to (American) English of roughly 6 years in a school setting. The second group consisted of 27 early bilingual listeners (11 male, 16 female) with Azerbaijani and Persian as their first two languages (see above for more explanation on their language background). The bilinguals were comparable to the monolinguals in all relevant aspects (age, exposure to English, level of education). They were tested in secondary schools in the city of Marand in the East Azerbaijan Province in the North West of Iran.

All listeners filled in the Language Experience and Proficiency questionnaire (LEAP-Q) developed by Marian et al. (2007). This questionnaire asks the participant to estimate their experience with and exposure to the languages they command, and to self-rate their proficiency and (non-)nativeness in each of these languages.

3.3 Procedure

Participants sat at a table in a small-size quiet office. They listened over a good quality headset (Sennheiser PC3) to the stimuli being played to them from a notebook computer. The monolingual participants saw a screen as shown in Figure 2A. On the screen, six common Persian words were printed in Arabic script (as is standard in the Iranian school system), which were identical in all respects except for the vowel. Each response button therefore corresponded to one of the six vowels of Persian, arranged in two columns (left: front vowels, right: back vowels) by three rows (top: high vowels, middle: mid vowels, bottom: low vowels). Participants

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

were instructed to listen to each English word being played just once, and to decide which of the six vowels contained in the words written in the response buttons resembled the vowel they had just heard most. Participants responded by clicking the button of their choice with a mouse pointer. Immediately after the mouse click the goodness scale at the bottom of the screen turned from grey to yellow. Participants then decided whether the vowel they had selected was a poor or a good token of its category on a 5-point scale (where 5 signified ‘good’). Then the screen was reset, and after 2 seconds the next stimulus word was made audible.

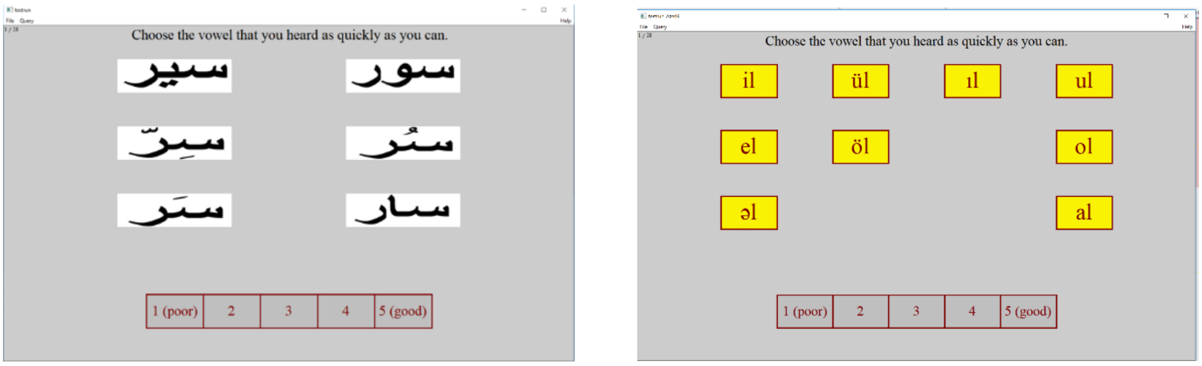

Figure 2. Screens showing the six response categories (in Arabic script) for the Persian version of the perceptual assimilation test (panel A, left). Panel B (right) shows the screen used for the Azerbaijani version of the test, with nine response categories in Azerbaijani orthography. The goodness scale at the bottom of the screen was activated only after the participant clicked a response vowel.

The 22 stimulus types were presented to each listener twice, in different quasi-random orders excluding immediate repetition of the same stimulus. A counter in the top-left corner of the screen kept the participants informed of their progress. No other feedback was given. Stimulus presentation and response collection was done through a Praat MFC script.

The bilingual participants took the assimilation test twice in immediate succession on the same day, the first time the Persian version (Figure 2A) and the second time in Azerbaijani. In the latter version, they saw the screen shown in Figure 2B, which contained nine rather than six vowel response buttons, with the keywords printed in standard Azerbaijani spelling. All instructions were in the target language, i.e., English. Again, the keywords were minimally different short words which differed in the vowel only. The words were arranged in four columns, corresponding, respectively, to front spread, front rounded, back spread and back rounded vowels. As for Persian, high, mid and low vowel buttons were listed on the top, middle and bottom row, respectively. The Azerbaijani spelling resembles that of Turkish, with the addition of the symbol ‘ə’, to represent open front [æ].

4 Results

Table 1A presents the perceptual assimilation results for the early monolingual Persian listeners. The 11 American English stimulus vowels are in the rows, while the six vowels of Persian they could be matched with are in the columns. Given 22 Persian listeners and 4 tokens of the same vowel (produced by two different male speakers), the maximum number of votes

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

for one particular response would be 88. The counts in the cells of the table have been converted to percentages in order to facilitate comparison with the results obtained by the 27 early bilingual listeners (yielding a maximum of 108 responses) in Tables 1B-C. Three columns in Table 1A-B have been greyed out; these are response vowels that occur in Azerbaijani only. Green cells in the table contain modal responses with ≥50% prevalence among the listeners. Yellow cells have responses on which between 25 and 50% of the listeners converged. When <25% prevalence was obtained for a response category, the cell has been left white. The left-hand smaller-sized number in each cell is the mean of the goodness ratings given to the particular stimulus-response pair. Next to it, we specify the so-called Fit-index, which is defined as the prevalence of the response category (expressed as a proportion) multiplied by the mean goodness rating (Guion et al., 2000).

Using simple decision rules, any non-native vowel that is assimilated to a native category in ≥50% of the responses, is considered *Categorized* (C). When there are two (or more) response categories with ≥25% of the votes, thereby adding up to ≥50% of the responses, the non-native vowel is considered *Uncategorized* (U). In all other cases, the non-native vowel is considered *Non-assimilable* (N). An alternative, more complicated, method categorizes the stimulus vowel as a Good, Fair, or Poor token of an L1 vowel depending on the distribution of the Fit-index over the modal responses per AE vowel (Guion et al., 2000). Following Wang and Chen (2019, 2020), when ≥1 SD above the mean of the modal Fit-indexes, the AE vowel is a Good token of the L1 vowel (G, green), between the mean and +1 SD it is a Fair token (F, yellow), between the mean and –1 SD it is a Poor token (P, orange), and any Fit-index <1 SD leaves the stimulus vowel unclassified (U, grey). Table 1 presents the results in terms of both methods.

When we apply the simple decision rules to the results in Table 1, we observe that ten out of eleven American English vowels qualify as C; the exception is the vowel [ʌ] (as in *but*), which is not assimilated to any of the vowels of either Persian or Azerbaijani by the bilingual listeners (i.e., N), whether responding in the Persian mode or in the Azerbaijani mode. Interestingly, [ʌ] is U in the perception of the monolingual Persian listeners, where the responses are equally divided between Persian /ɑ/ and /o/. Given that [ʌ] is a central vowel, it is not surprising that the three central vowels in Azerbaijani compete with /ɑ/ and /o/ and together draw 42% of the responses, thereby depleting other response categories. When the bilinguals respond in the Persian mode, only /o/ is chosen in >25% of the responses, without reaching the majority criterion for C. Here the bilinguals differ from the monolinguals even when they respond in the same mode, i.e., Persian.

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

Monolingual Persian (N = 22) — Mean = 3.0; SD = 1.01

Stimulus vowel			i	y	e	œ	æ	ɑ	o	ɯ	u	Fit-index	category
heed	i	C	87 4.3 3.7		13 3.5 .5							3.7	G
hid	ɪ	C	39 3.9 1.5		60 3.9 2.3						1 3.0 .0	2.3	P
hayed	e	C	39 3.6 1.4		61 3.5 2.1							2.1	P
head	ɛ	C	2 4.0 .1		90 4.0 3.6		6 3.8 .2	1 4.0 .0				3.6	F
had	æ	C	2 4.5 .1		37 4.1 1.5		61 4.4 2.7					2.7	P
hud	ʌ	U			7 2.8 .2		1 4.0 .0	37 4.0 1.5	39 4.2 1.6		15 4.1 .6	1.6	U
hod	ɑ	C					4 3.3 .1	96 4.5 4.3				4.3	G
hawed	ɔ	C						95 4.4 4.2	5 4.5 0.2			4.2	G
hoed	o	C						1 3.0 .0	52 3.8 2.0		46 3.9 1.8	2.0	P
hood	ʊ	C			1 3.0 .0		2 4.5 .1	1 4.0 .0	56 4.2 2.4		39 4.1 1.6	2.4	P
who'd	u	C							19 3.8 0.7		81 4.3 3.5	3.5	F

B. Bilingual Persian (N = 27) — Mean = 2.9; SD = 0.82

Stimulus vowel			i	y	e	œ	æ	ɑ	o	ɯ	u	Fit-index	category
heed	i	C	98 4.0 3.9		2 2.5 0.1							3.9	G
hid	ɪ	C	58 3.7 2.1		38 3.5 1.3				3 3.0		1 4.0 .0	2.1	P
hayed	e	C	25 3.7 .9		74 3.6 2.7			1 1.0 .0				2.7	P
head	ɛ	C	5 4.0 .2		93 3.7 3.4		2 3.5 .1	1 3.0 .0				3.4	F
had	æ	C	2 2.5 .1		41 3.8 1.6		56 4.0 2.2		1 4.0 .0			2.2	P
hud	ʌ	N	2 4.0 .1		9 2.9 .3		1 2.0 .0	20 3.8 .8	48 3.8 1.8		19 3.3 .6	1.8	U
hod	ɑ	C	1 4.0 .0				2 3.5 .1	95 4.1 3.9	2 4.0 .1			3.9	G
hawed	ɔ	C			2 2.5 .1		1 3.0 .0	84 4.2 3.5	11 3.5 .4		2 3.5 .1	3.5	F
hoed	o	C	2 4.0 .1		1 4.0 .0				61 4.0 2.4		36 3.8 1.4	2.4	P
hood	ʊ	C	2 4.0 .1		3 4.0 .1			3 3.3 .1	41 3.9 1.6		52 3.9 2.0	2.0	U
who'd	u	C	1 4.0 .0					1 1.0 .0	7 3.8 .3		91 4.1 3.7	3.7	G

C. Bilingual Azerbaijani (N = 27) — Mean = 2.2; SD = 0.96

Stimulus vowel			i	y	e	œ	æ	ɑ	o	ɯ	u	Fit-index	category
heed	i	C	75 3.8 2.9		11 3.6 .4					14 3.8 05		2.9	F
hid	ɪ	C	38 3.7 1.4	2 3.5 .1	36 3.5 1.3	3 3.0 .1	2 2.0 .0			18 3.8 .7	2 4.0 .1	1.4	P
hayed	e	C	7 2.9 .2		77 3.8 2.9	2 2.5 .1	7 3.4 .2	1 4.0 .0		5 3.6 .2	1 5.0 .1	2.9	F
head	ɛ	C	6 3.1 .2	1 3.0 .0	77 3.8 2.9	1 4.0 .0	8 3.8 .3			6 3.2 .2	1 4.0 .0	2.9	F
had	æ	C	2 3.0 .1		22 3.4 .7	1 3.0 .0	65 3.8 2.5	6 4.1 .2	1 4.0 .0	3 2.7 .1		2.5	F
hud	ʌ	N	2 2.0 .0	17 3.7 .6	6 3.0 .2	16 3.6 .6	4 3.0 .1	17 3.6 .6	18 3.5 .6	9 3.3 .3	12 3.7 .4	0.6	U
hod	ɑ	C	1 3.0 .0			1 1.0 .0	9 3.0 .3	83 4.0 3.3	1 4.0 .0	3 4.3 .1	2 2.5 .1	3.3	G
hawed	ɔ	C		1 4.0 .0		1 4.0 .0	6 3.2 .2	83 3.9 3.2	1 4.0 .0	5 4.4 02	4 2.8 .1	3.2	G
hoed	o	C		15 3.7 .6		48 3.8 1.8			28 3.8 1.1	2 2.0 .0	7 3.8 .3	1.8	P
hood	ʊ	C	1 4.0 .0	24 3.7 .9	6 3.0 .2	30 3.4 1.0	2 3.5 .1	1 3.0 .0	15 3.6 .5	4 3.0 .1	19 3.8 .7	1.0	U
who'd	u	C		38 3.8 1.4	1 5.0 .1	13 3.4 .4			24 3.8 .9	1 4.0 .0	23 3.8 .9	1.4	P

Table 1. Perceptual assimilation of eleven vowels of American English to the six vowels of Persian by early monolingual Persian listeners (A), and by early bilingual Azerbaijani-Persian listeners (B). Panel C shows the results of the bilinguals when instructed to assimilate the English vowels to the nine vowels of Azerbaijani. The three added vowel response categories have been greyed out in the Persian response mode. The prevalence of a response vowel is expressed as a percentage (large print). Green cells contain responses with ≥50% agreement. Yellow cells contain responses with 25 to 50% agreement. The numbers in small print are the mean goodness ratings for the particular stimulus-response pair (left) and the Fit-index (right, bolded). Categorization based on Fit-index is indicated in the right-hand margin. Here G = Good, F = Fair, P = Poor, U = unclassified (for details see text).

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

Given an inventory of eleven, there are (11 × 10) / 2 = 55 possible contrasts between American English monophthongs. The majority of these pairwise contrasts are between vowels that are non-adjacent in the English vowel space, and which assimilate to different vowel categories in either Persian or in Azerbaijani. For instance, in Table 1A, English /i/ is assimilated to Persian /i/ in 87% of the judgments with a goodness rating of 4.3 and a Fit-index of 3.7 (Good). English /u/ assimilates to Persian /u/ with 81% with a goodness of 4.3 and Fit-index of 3.5 (Fair). This is an example of a Two Category (TC) assimilation scenario, for which the prediction is that the members of the contrast are easily discriminated by Persian learners of English.

In all, there are 40 TC pairs in Table 1A. Of the remaining 15 contrasts, ten involve the U vowel /ʌ/, which is paired with one of the other ten vowels, all of which are C. Since one vowel in the UC pair maps onto a known category in the L1 while the other vowel remains a difficult choice, the members of the pair should be rather easy to discriminate. The prediction, therefore, is that the Persian learners will quickly realize that the vowel /ʌ/ is different from any vowel in their own language – so that they are prompted to set up a new category for the non-native sound.

Five more contrasts are between vowels that are adjacent in the English vowel space. Three of these are in a Single Category scenario: both /ɪ, e/ assimilate to Persian /e/ (Fit-indexes: 2.3, 2.1 = Poor-Poor), both /ɑ, ɔ/ map onto Persian /ɑ/ (Fit-indexes: 4.3, 4.2 = Good-Good), and both /o, ʊ/ assimilate to Persian /o/ (Fit-indexes: 2.0, 2.4 = Poor-Poor). Learners will find the SC members difficult to distinguish, and will be challenged to learn how to split their single native category into two new, smaller categories. Without explicit instruction, the learner will never be aware that the contrast exists in the foreign language and will conflate the categories forever (Flege, 1995).

Finally, two contrasts in Table 1A are between adjacent English vowels, one of which is a convincing token of a Persian vowel, while the other, although it assimilates to the same category, is a clearly poorer exemplar of it. This Category Goodness (CG) scenario applies to the contrasts /ɪ, ɛ/ and /e, ɛ/. All three vowels map onto Persian /e/, but /ɛ/ is perceived as a much better exemplar (Fit-index: 3.6 = Fair) of Persian /e/ than either /ɪ/ (too high, Fit-index: 2.3 = Poor) or /e/ (too long and diphthongal, Fit-index: 2.1 = Poor) is. Fairly good discrimination is predicted for CG pairs, as it will be easier for the learner to discover that and how his/her single native category has to be subdivided into one part that matches with exemplars rather close to the Persian prototype, while the other part is more remote from the prototype and contains non-typical exemplars only.

When the bilinguals respond in the Persian mode, 41 of the 55 pairwise contrasts are of the TC type. The mid-low central vowel /ʌ/ is N, i.e., is outside the Persian vowel space – but only just. The lower central region of the vowel space is unused in Persian. As a result, all ten contrasts involving /ʌ/ are of the NC type – for which scenario the same prediction holds as for the UC type with the monolingual EFL learners. There are no CG contrasts in Table 1B, but four contrasts are of the SC type: /i, ɪ/, which both assimilate to Persian /i/ (Fit-indexes: 3.9, 2.1 = Good-Poor), /e, ɛ/ assimilate to /e/ (Fit-indexes: 2.7, 3.4 = Poor-Fair), /ɑ, ɔ/ assimilate to /ɑ/ (Fit-indexes: 3.9, 3.5 = Good-Fair), and /u, ʊ/ assimilate to /u/ (Fit indexes: 3.7, 2.0 = Good-Uncategorized).

In Table 1C, where the bilinguals respond in the Azerbaijani mode, the number of native vowel categories is nine rather than six. Since there are more response categories, it is harder for a non-native vowel to meet the 25% or 50% lower-bound criteria. This shows up as a smaller number of non-native vowels that are C: the vowels /ɪ/ and /o/ are U, while /ʌ/, /u/ and /ʊ/ are

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

N. This reduces the number of easy TC contrasts to 13. Only two contrasts are of the difficult SC type, i.e., /e, ɛ/ both assimilate to Azerbaijani /e/ (Fit-indexes: 2.9, 2.9 = Fair-Fair), and /ɑ, ɔ/ to Azerbaijani /ɑ/ (Fit-indexes: 3.3, 3.2 = Good-Good). The remaining 40 contrasts are classified as UC (14), NC (20), NU (5) or NN (1). Interestingly, there are no instances of the CG scenario. It should be noted that the classification of several of these contrasts into scenarios may be unduly rigid. For instance, /ʊ/ would have been counted as U (between /œ/ and /y/) with 1 point more /y/-responses; /u/ would have been U (between /y/ and /o/) with 1 point more /o/-responses. Similarly, /o/ would have been C instead of U with 2 points more /œ/-responses.

Of the three additional, central vowels in Azerbaijani, two are close enough to the AE stimulus vowels to attract majority responses in excess of 25%. The vowel /y/ is chosen as the nearest equivalent of AE /u/ with a Fit-index of 1.4, which makes it a poor but recognizable token of /y/, rather than of /o/ or /u/. This shows that the bilingual listeners are sensitive to the fronting of AE /u/, so that it should be easier for them to set up an adequate perceptual representation of AE /u/, and achieve a more authentic pronunciation of this vowel.

In similar vein, the Azerbaijani central /œ/ attracts majority responses in excess of 25% for both AE /o/ (48%, Fit-index: 1.8) and for /ʊ/ (30%, Fit-index: 1.0). Therefore, the /o, ʊ/ pair constitutes a PU contrast, which should not constitute a great learning problem (comparable to a CG scenario). The attractiveness of /œ/ also shows that the bilinguals perceive the more centralized vowel quality of short /ʊ/ and of (the onset of) /o/.

Overall, the Fit-indexes for the modal responses per stimulus vowel are lower when the assimilation task involves a choice from nine response alternatives (Azerbaijani) than when the same (early bilingual) listeners respond in the Persian mode with six response categories (means: 2.87 vs. 2.27). Monolingual listeners have slightly higher Fit-indexes than the bilinguals when both respond in Persian mode: 3.00 vs. 2.87). The overall effect is significant by a one-way Repeated-Measures Analysis of Variance, $F(2, 20) = 10.4$ ($p = .001$, $p\eta^2 = .510$). Bonferroni post-hoc tests ($\alpha = .050$) confirm that the Azerbaijani Fit-indexes are significantly lower than those for the two Persian response sets, which do not differ significantly from each other. Moreover, the Fit-indexes for the monolinguals and bilinguals (responding in Persian mode) are very strongly correlated ($r = .917$, $p < .001$). The correlations are appreciably lower when the Azerbaijani response set is compared ($r = .700$, $p = .017$ with monolinguals, $r = .693$, $p = .018$ with bilinguals).

Table 2 summarizes the most important differences in the perceptual assimilation patterns observed between monolinguals and bilinguals in Persian response mode, as well as between the bilinguals in Persian versus Azerbaijani mode. The table exemplifies that the monolinguals have different assimilation patterns than the bilinguals, even when the latter respond in Persian mode. It suggests that the major difficulties will be in the contrast between the short and long high-mid vowels (both front and back) for the monolinguals, while keepinng the high-mid vowels separate from the high vowels as well as from the (mid-)low vowels. The bilinguals, however, are predicted to experience difficulties in distinguishing the high vowels from the mid-high vowel on the one hand, and the mid-high from the mid-low vowels on the other hand. This also suggests that the bilinguals attend to differences in vowel quality and not so much to differences in vowel duration. The bilinguals' assimilation patterns for for AE front vowels are the same, whether they respond in Persian mode or in Azerbaijani mode. For the back vowels, the Azerbaijani response mode generates a different pattern due to the attractiveness of two of the central vowels (see above).

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

<table>
<tr><th rowspan="2"></th><th>Monol.</th><th rowspan="2">AE</th><th colspan="2">Bilinguals</th><th rowspan="2">AE</th><th rowspan="2"></th></tr>
<tr><th>Persian</th><th>Persian</th><th>Azerb.</th></tr>
<tr><td rowspan="4">Front vowels</td><td></td><td>heed</td><td rowspan="2">/i/</td><td rowspan="2">/i/</td><td>heed</td><td>i</td></tr>
<tr><td rowspan="2">/e/</td><td>hid</td><td>hid</td><td>ɪ</td></tr>
<tr><td>hayed</td><td rowspan="2">/e/</td><td rowspan="2">/e/</td><td>hayed</td><td>e</td></tr>
<tr><td></td><td>head</td><td>head</td><td>ɛ</td></tr>
<tr><td rowspan="5">Back vowels</td><td></td><td>who’d</td><td rowspan="2">/u/</td><td></td><td>who’d</td><td>u</td></tr>
<tr><td rowspan="2">/o/</td><td>hood</td><td rowspan="2">/œ/</td><td>hood</td><td>ʊ</td></tr>
<tr><td>hoed</td><td></td><td>hoed</td><td>o</td></tr>
<tr><td rowspan="2">/ɑ/</td><td>hawed</td><td rowspan="2">/ɑ/</td><td rowspan="2">/ɑ/</td><td>hawed</td><td>ɔ</td></tr>
<tr><td>hod</td><td>hod</td><td>ɑ</td></tr>
</table>

Table 2. Summary of Same Category (SC) contrasts in American English (AE) vowels as perceived by monolingual Persian EFL learners (left) and by bilingual EFL learners in either Persian (L2) or Azerbaijani (L1) mode. Two AE vowels joined by a brace denote a SC contrast (reddish cells); a yellow cell denotes a Category Goodness (CG) contrast or similar.

The low-back AE vowels /ɔ, ɑ/, both of which are pronounced long (see Appendix, Figure A2), project onto the same vowel in Persian as well as in Azerbaijani for both groups of EFL learners. Although this predicts a learning problem for the EFL learners, the general advice is not to make this a priority in the teaching of American English pronunciation. The /ɔ, ɑ/ contrast has a low functional load, and is subject to the low-back vowel merger in the pronunciation of most native speakers of American English, who then pronounce all low back vowels as low /ɑ/ (e.g., Ladefoged & Johnson, 2011: 212–213 ; Carley & Mees, 2020).

5 Conclusions and discussion

The first question we asked was how the two groups of EFL learners assimilate the vowels of English to those of Persian, and – in the case of the early bilinguals – to those of Azerbaijani. The results bear out that of the eleven vowels of American English the great majority (i.e., ten) are Categorized in Persian, by both the monolingual and the bilingual EFL learners. The only Uncategorized vowel is /ʌ/, which makes sense because Persian has no central vowels. The critical SC contrast scenario is predicted for Persian EFL learners in only three vowel pairs, all of which involve members that are adjacent in the American English vowel quality space. These are the AE vowel pairs /ɪ, e/, /ɑ, ɔ/ and /o, ʊ/. One SC contrast, between the half-open back vowels /ɑ, ɔ/, can be ignored as a learning problem, because this contrast is also absent in most varieties of American English (see above).

Neither Persian nor Azerbaijani has a tense-lax vowel contrast. This is reflected by the SC assimilation pattern observed for /e, ɪ/ and /o, ʊ/ for the monolinguals, and in /i, ɪ/ and /u, ʊ/ for the bilinguals. The bilinguals have another SC pair in /e, ɛ/, which would suggest that their Persian category /e/ is larger than its counterpart is for the monolinguals – while the bilinguals’ category for /i/ would be smaller, possibly due to competition from Azerbaijani /y/. The assimilation of the tense-lax pairs in the high/mid-section of the vowel space would be the most important difference between the monolingual and bilingual participants. This may be due to some form of interaction on the part of the bilinguals with the competing vowel system of

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

Azerbaijani. Azerbaijani has three central vowels in the high-mid part of the vowel space, so that the dispersion area for the peripheral vowels /i, e, o, u/ is more limited in the bilingual Persian vowel system than for monolingual Persians.

The predictions derived here from the perceptual assimilation results can be provisionally checked against literature data on the acoustic and perceptual discrimination of the British English (SSBE) vowels by monolingual Azerbaijani and Persian EFL learners. The data show that both groups of EFL learners experience the same problems in their production and perceptual discrimination of the English monophthongs. Azerbaijani EFL learners showed poor perceptual discrimination in four SSBE vowel pairs: /ʌ~ɒ/, /ɑː~ʌ/, /uː~ʊ/, /ɑː~ɒ/ (.45 < A' <.58) but not for the other seven pairs tested: /ɔː~ɒ/, /æ~ʌ/, /iː~ɪ/, /e~æ/, /ɔː~ʊ/, /iː~e/, /ɑː~ɔː/ (.75 < A' < .87).[1] No comparison is available with SSBE native listeners. The differences in perceptual discrimination were echoed in the formant structure and/or duration of the vowel production by the same learners (Ghaffarvand Mokari & Werner, 2017). No data on perceptual discrimination of SSBE vowel pairs by monolingual Persian EFL learners are available at this time. However, insufficient acoustic contrast was observed for the pairs /ʌ~ɒ/, /iː~ɪ/, /uː~ʊ/ but not for /e~æ/ (Pillai & Delavari, 2012). This suggests that the three extra central vowels in the inventory of Azerbaijani do not provide an advantage for Azerbaijani over Persian EFL learners, which parallels the predictions derived from our perceptual assimilation results for American English vowels.

Testing the predictions derived from the present PAM study will also allow us to check whether the categorization of the AE vowel contrasts based on simple decision rules should be refined by including the differences in Fit-indexes of the members of contrasts, on the basis of which some of the SC contrasts may have to be regarded as CG contrasts – with different degrees of fit between the members.

More specific tests of the predictions made on the basis of our PAM test will be reported in the next stage of our project, by mapping out the perceptual representation of the American English vowels by our learners and comparing it to that of American native listeners (see Van Heuven et al., 2020 for preliminary results), and by studying the acoustic characteristics of the AE vowels produced by our learners and by American native speakers.

Provisonal analysis of both the perceptual representation data and the EFL learners’ AE vowel production in Afshar’s (2022) dissertation, as well as informal listening to the EFL recordings, reveals that the Iranian EFL learners have not grasped all perceptually relevant the details of the American English (monophthongal) vowels.[2] Afshar’s results show that the learners tend to distinguish rather adequately between AE short and long vowels, so that correlation of vowel duration between American native speakers and the EFL learners is substantial, even though all the vowels are produced much shorter in EFL than in L1 American speech, most likely because the learners fail to properly lengthen the vowels in the context before voiced coda /d/. It is known from the literature that differences in vowel duration are easily noticed by foreign-language learners, even if the learners’ native language has no length contrasts (e.g., Van Heuven 1986; Bohn 1995; Escudero 2005; Perwitasari 2019).

[1] Ghaffarvand Mokari and Werner (2017) adopted the non-parametric A' (A prime) coefficient as their measure of discrimination accuracy. Values near 1.0 indicate good discrimination while .5 means chance performance. A' = .5 + (((H–F) (1+H–F)) / (4H(1–F))) when H (proportion of Hits) ≥ F (proportion of False Alarms) and .5 – (((H–F) (1+H–F)) / (4H(1–F))) when H < F (e.g., Stanislaw & Todorov, 1999: 142).

[2] A draft version of this dissertation was defended in April 2022. The revision, based on comments received, will be publicly available later this year.

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

However, the contrasts in vowel quality (spectral contrasts) between all neighboring vowels in the American English vowel space, with the exception of pairs involving /æ/, are severely blurred, for both groups of EFL learners tested. Moreover, the confusion structure in the results of the monolingual and early bilingual EFL learners appears to be highly similar, suggesting that transfer effects are dominated by the Persian vowel system, even for the early bilinguals. It would seem, therefore, that there is much room for improvement for the EFL pronunciation instruction in the Iranian secondary-school system, at least as far as the monopthongal vowels are concerned.

References

Afshar, N. (2022): *Production and perceptual representation of American English vowel sounds by monolingual Persian and early bilingual Azerbaijani-Persian adolescents*. Doctoral dissertation. Veszprém: Multilingualism Doctoral School, University of Pannonia.

Ansarin, A. A. (2004): An acoustic analysis of Modern Persian vowels. *Proceedings of the 9th Conference on Speech and Computer (SPECOM), September 20–22, 2004, St. Petersburg*. https://www.isca-speech.org/archive_open/specom_04/spc4_315.pdf (Accessed 20 May 2022)

Aronov, R., McHugh, B. D. & Molnar, T. (2017): A pilot acoustic study of Modern Persian vowels in colloquial speech. *Proceedings of the Linguistic Society of America* 2.1, 1–7. http://dx.doi.org/10.3765/plsa.v2i0.4059

Best, C. T. (1995): A direct realist perspective on cross-language speech perception. In: Strange, W. (ed.): *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*. Timonium, MD: York Press, 167–200.

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. *Journal of the Acoustical Society of America* 109.2, 775–794. http://dx.doi.org/10.1121/1.1332378

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* 4.3, 45–60. http://dx.doi.org/10.1037//0096-1523.14.3.345

Best, C. T. & Tyler, M. D (2007): Nonnative and second-language speech perception: Commonalities and complementarities. In: Munro, M. J. & Bohn, O.-S. (eds.): *Language experience in second language speech learning: In honor of James Emil Flege.* Amsterdam: John Benjamins, 13–34. https://www.academia.edu/24865551/ (Accessed 20 May 2022). https://doi.org/10.1075/lllt.17.07bes

Bigdeli, N. & Sadeghi, V. (2020): Perceptual evidence for the phonological adaptation of English vowels in Persian sound system. *Scientific Journal of Language Research* 12, 274–295. http://dx.doi.org/10.22051/jlr.2019.23487.1629

Boersma, P. & Van Heuven, V. J. (2001): Speak and unSpeak with Praat. *Glot International* 5, 9–10, 341–347. https://www.fon.hum.uva.nl/paul/papers/speakUnspeakPraat_glot2001.pdf (Accessed 20 May 2022)

Boersma, P. & Weenink, D. (2019): Praat, a system for doing phonetics by computer. www.praat.org (Accessed 20 May 2022)

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

Bohn, O.-S. (1995): Cross-language speech perception in adults: First language transfer doesn't tell it all. In: Strange, W. (ed.): *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, MD: York Press, 279–304.

Brown, A. (1991): Functional load and the teaching of pronunciation. In: Brown, A. (ed.): Teaching English pronunciation: A book of readings. London: Routledge, 221–224.

Carley, P. & Mees, I. M. (2020): *American English Phonetics and pronunciation practice*. Abingdon, New York: Routledge. https://doi.org/10.4324/9780429492228

Celce-Murcia, M., Brinton, D. M. & Goodwin, J. M. (1996): *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.

Cutler, A. (2012): *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/9012.001.0001

Eckman, F. R. (1977): Markedness and the contrastive analysis hypothesis. *Language Learning* 27.2, 315–330. http://dx.doi.org/10.1111/j.1467-1770.1977.tb00124.x

Eckman, F. R. (1985): Some theoretical and pedagogical implications of the markedness differential hypothesis. *Studies in Second Language Acquisition* 13.1, 23–41. https://doi.org/10.1017/S0272263100005544

Escudero, P. (2005): *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. LOT Dissertation Series nr. 113. LOT. https://www.lotpublications.nl/Documents/113_fulltext.pdf (Accessed 20 May 2022)

Flege, J. E. (1987): The production of 'new' and 'similar' phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15.1, 47–65. http://dx.doi.org/10.1016/S0095-4470(19)30537-6

Flege, J. E. (1995): Second language speech learning: theory, findings, and problems. In: Strange, W. (ed.): *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*. Timonium, MD: York Press, 233–277.

Flege, J. E. & Bohn, O.-S. (2020): The revised Speech Learning Model (SLM-r). In: Wayland, R. (ed.): *Second language speech learning: Theoretical and empirical progress*. Cambridge: Cambridge University Press, 3–83. http://dx.doi.org/10.1017/9781108886901.002

Frieda, E. M., Walley, A. C., Flege, J. E. & Sloane, M. E. (1999): Adults' perception of native and nonnative vowels: implications for the perceptual magnet effect. *Perception and Psychophysics* 61.3, 561–577. http://dx.doi.org/10.3758/bf03211973

Ghaffarvand Mokari, P. & Werner, S. (2016): An acoustic description of spectral and temporal characteristics of Azerbaijani vowels. *Poznan Studies in Contemporary Linguistics* 52.3, 503–518. http://dx.doi.org/10.1515/psicl-2016-0019

Ghaffarvand Mokari, P. & Werner, S. (2017): Perceptual assimilation predicts acquisition of foreign language sounds: The case of Azerbaijani learners' production and perception of Standard Southern British English vowels. *Lingua* 185, 181–195. http://dx.doi.org/10.1016/j.lingua.2016.07.008

Ghaffarvand Mokari, P., Werner, S. & Talebi, A. (2017): An acoustic description of Farsi vowels produced by native speakers of Tehrani dialect. *The Phonetician* 114, 6–23. https://www.researchgate.net/profile/Payam-Ghaffarvand-Mokari/publication/319261459 (Accessed 20 May 2022)

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

Guion, S. G., Flege, J. E., Akahane-Yamada, R. & Pruitt, J. C. (2000): An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* 107.5, 2711–2724. http://dx.doi.org/10.1121/1.428657

Hillenbrand, J., Getty, L. A., Clark, M. J. & Wheeler, K. (1995): Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97.5, 3099–3111. http://dx.doi.org/10.1121/1.411872

Howlader, M. R. (2010): Teaching English pronunciation in countries where English is a second language: Bangladesh perspective. *ASA University Review* 4.2, 233–244. www.asaub.edu.bd/data/asaubreview/v4n2sl20.pdf (Accessed 20 May 2022)

Jenkins, J. (2000): *The phonology of English as an international language: New models, new norms, new goals*. Oxford: Oxford University Press.

Jenkins, J. (2002): *World Englishes*. London, New York: Routledge.

Jones, T. (2019): A corpus phonetic study of contemporary Persian vowels in casual speech. *University of Pennsylvania Working Papers in Linguistics* 25, article 15. https://repository.upenn.edu/pwpl/vol25/iss1/15 (Accessed 20 May 2022)

Kuhl, P., Conboy, B., 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 B: Biological Sciences* 363, 979–1000. http://dx.doi.org/10.1098/rstb.2007.2154

Mirahadi, S. S., Kamran, F., Mansuri, B., Tohidast, S. A., Rashtbari, K., Panahgholi, E. & Tagipur, A. (2018): Investigation of the formant structure of Persian vowels in Persian-Azari bilingual adults. *Archives of Rehabilitation* 19, 142–148. https://doi.org/10.32598/rj.19.2.142

Munro, M. J. & Derwing, T. M. (2006): The functional load principle in ESL pronunciation instruction: An exploratory study. *System* 34.4, 520–531. http://dx.doi.org/10.1016/j.system.2006.09.004

Ladefoged, P. & Disner, S. F. (2012): *Vowels and consonants*. Chichester: Wiley.

Ladefoged, P. & Johnson, K. (2011): *A course in Phonetics*. Boston, MA: Wadsworth.

Lado, R. (1957): Languages across cultures. Applied Linguistics for language teachers. Ann Arbor, MI: University of Michigan Press.

Majidi, M.-R., & Ternes, E. (1999): Persian (Farsi). *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 124–125. https://archive.org/details/rosettaproject_pes_phon-2 (Accessed 20 May 2022)

Mannell, R., Cox, F., & Harrington, J. (2009): An introduction to Phonetics and Phonology, Macquarie University. http://clas.mq.edu.au/speech/phonetics/ (Accessed 20 May 2022)

Marian, V., Blumenfeld, H. K. & Kaushanskaya, M. (2007): The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research* 50, 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Morley, J. (1991): The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly* 25.3, 481–521. http://dx.doi.org/10.2307/3586981

Munro, M. J. & Derwing, T. M. (2006): The functional load principle in ESL pronunciation instruction: An exploratory study. *System* 34, 520–531. http://dx.doi.org/10.1016/j.system.2006.09.004

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

Peivasti, S. M. (2012): An acoustic analysis of Azerbaijani vowels in Tabrizi dialect. *Journal of Basic and Applied Scientific Research* 2, 7181–7184. https://www.textroad.com/pdf/JBASR/J.%20Basic.%20Appl.%20Sci.%20Res.,%202(7)7181-7184,%202012.pdf (Accessed 20 May 2022)

Perwitasari, A. (2019): *The acquisition of English vowels by Javanese and Sundanese native speakers* (LOT Dissertation Series 532). Utrecht: LOT. https://www.lotpublications.nl/Documents/532_fulltext.pdf (Accessed 20 May 2022)

Peterson, G. E. & Barney, H. L. (1952): Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24, 175–184. http://dx.doi.org/10.1121/1.1906875

Pillai, S. & Delavari, H. (2012): The production of English monophthong vowels by Iranian EFL learners. *Poznań Studies in Contemporary Linguistics* 48, 473–493. http://dx.doi.org/10.1515/psicl-2012-0022

Sadeghi, V. & Bigdeli, N. (2018): انطباق واجی واکه های انگلیسی با فارسی در چارچوب انگارهٔ شنیداری [Phonological adaptation of English vowel with Persian vowel based on perceptual assimilation]. *Journal of Research in Linguistics* 10, 43–60. https://www.sid.ir/FileServer/JF/35413971803.pdf (Accessed 20 May 2022)

Salehi, M. & Neysani, A. (2017): Receptive intelligibility of Turkish to Iranian-Azerbaijani speakers. *Cogent Education* 4.1, 1–15. https://doi.org/10.1080/2331186X.2017.1326653

Stanislaw, H. & Todorov, N. (1999): Calculation of signal detection theory measures. *Behavior Research, Methods, Instruments, & Computers* 31.1, 137–149. http://dx.doi.org/10.3758/BF03207704

Strange, W., Bohn, O.-S., Trent, S. A. & Nishi, K. (2004): Acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America* 115, 1791–1807. http://dx.doi.org/10.1121/1.1687832

Traunmüller, H. (1990): Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America* 88, 97–100. http://dx.doi.org/10.1121/1.399849

Trofimovich, P. & Baker, W. (2006): Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition* 28, 1–30. http://dx.doi.org/10.1017/S0272263106060013

Tyler, M. D. (2019): PAM-L2 and phonological category acquisition in the foreign language classroom. In: Nyvad, A. M., Hejná, M., Højen, A., Jespersen A. B. & Sørensen M. H. (eds.): *A sound approach to language matters – In honor of Ocke-Schwen Bohn*. Aarhus: Dept. of English, School of Communication & Culture, Aarhus University, 607–630). https://www.researchgate.net/publication/333124848 (Accessed 20 May 2022)

Van Heuven, V. J. (1986): Some acoustic characteristics and perceptual consequences of foreign accent in Dutch spoken by Turkish immigrant workers. In: Van Oosten, J. & Snapper, J. F. (eds.): *Dutch Linguistics at Berkeley, papers presented at the Dutch Linguistics Colloquium held at the University of California, Berkeley on November 9th, 1985*. Berkeley, CA: The Dutch Studies Program, U.C. Berkeley, 67-84. https://openaccess.leidenuniv.nl/handle/1887/2831 (Accessed 20 May 2022)

Van Heuven, V. J., Afshar, N., & Disner, S. F. (2020): Mapping perceptual vowel spaces in native and foreign language: Persian learners of English compared with American native speakers. In: Bátyi, Sz. & Lengyel, Zs. (eds.): *Magyar és nem magyar kontextus. Tanulmányok Navracsics Judit köszöntésére/Bilingualism: Hungarian and non-Hungarian context. Studies in honor of Judit Navracsics*. Veszprém: Pannon Egyetem, 113–130. https://www.researchgate.net/publication/347252430 (Accessed 20 May 2022)

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

Van Heuven, V. J. & Kirsner, R. S. (2004): Phonetic or phonological contrasts in Dutch boundary tones? In: Cornips, L. & Doetjes, J. (eds.): *Linguistics in the Netherlands 2004*. Amsterdam: John Benjamins, 102–113. http://dx.doi.org/10.1075/avt.21.13heu

Walker, R. (2001): Pronunciation for international intelligibility. *English Teaching Professional* 21, 1–7. https://englishglobalcom.files.wordpress.com/2013/12/pronunciation-for-international-intelligibility2.pdf (Accessed 20 May 2022)

Wang, H., & Van Heuven, V. J. (2006): Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers. In: Van de Weijer, J. M. & Los, B. (eds.): *Linguistics in the Netherlands 2006.* Amsterdam: John Benjamins, 237–248. http://dx.doi.org/10.1075/avt.23.23wan

Wang, H. & Van Heuven, V. J. (2014): Is a shared interlanguage beneficial? Mutual intelligibility of American, Dutch and Mandarin speakers of English. In: Van den Doel, R. & Rupp, L. (eds.): *Pronunciation Matters. Accents of English in The Netherlands and elsewhere*. Amsterdam: VU University Press, 175–194. https://docplayer.net/34448641-Pronunciation-matters.html#download_tab_content (Accessed 20 May 2022)

Wang, X. & Chen, J. (2019): English speakers' perception of Mandarin consonants: The effect of phonetic distances and L2 experience. In: Calhoun, S., Escudero, P., Tabain, M. & Warren, P. (eds.): *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia* Australasian Speech Science and Technology Association, 250–254. https://assta.org/proceedings/ICPhS2019/papers/ICPhS_299.pdf (Accessed 20 May 2022)

Wang, X. & Chen, J. (2020): The Acquisition of Mandarin consonants by English learners: The relationship between perception and production. *Languages* 5.2, 1–15. http://dx.doi.org/10.3390/languages5020020

Naiemeh Afshar, MA
University of Pannonia
Multilingualism Doctoral School
Egyetem u. 10
H-8200 Veszprém
afshar.naeimeh@mftk.uni-pannon.hu

Prof. dr. Vincent J. van Heuven
University of Pannonia
Multilingualism Doctoral School
Egyetem u. 10
H-8200 Veszprém
v.j.j.p.van.heuven@hum.leidenuniv.nl

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

Appendix. Acoustic details of American English vowel tokens used in the PAM test.

The vowel tokens were segmented from the first clear glottal pulse following the [h]-noise until the near-silence of the [d] following the vowel. Mean formant frequencies F1 and F2 were measured (using the Burg algorithm in Praat with three formants in a 0 to 3 KHz frequency band), from the vowel onset until either F1 or F2 showed the beginning of a transition to the following [d]. For the semi-diphthongs /e/ and /o/, mean F1 and F2 were computed for the first 50% of the vowel duration only. The results are shown in Table A1.

Vowel	Speaker 1			Speaker 2		
	F1	F2	Dur	F1	F2	Dur
i	300	2378	220	296	2163	184
ɪ	399	1989	113	403	1778	136
e	382	2157	293	436	1890	240
ɛ	506	1887	184	524	1688	128
æ	649	1730	273	667	1637	216
u	352	1184	245	353	1212	205
ʊ	433	1339	209	465	1201	160
o	442	1047	289	424	1213	205
ɔ	550	967	278	589	873	268
ɑ	713	1150	249	620	992	258
ʌ	595	1154	172	523	1186	150

Table A1. Stimulus analysis of 22 vowel tokens used in PAM test. F1, F2 (Hz) and duration (ms) of eleven vowel tokens produced by two male American native speakers in /h..d/ context.

The formants were psychophysically scaled to Bark units so that equal distances in the F1-F2-plane correspond to equal auditory distances in vowel quality using the formula in Traunmüller (1990). The resulting vowel plot is shown in Figure A1. Note the similarity between the acoustic vowel chart here and the articulatory IPA diagram in Figure 1C.

A plot of the vowel durations is shown in Figure A2. There is a split in duration between the seven phonetically tense and long vowels and the four lax and short vowels. There are two vowel pairs in Figure A1 the members of which are spectrally close to one another. These are the pairs /ɪ, e/ and /ʊ, o/. These members will nevertheless be distinct by the difference in duration, and by the slight change in quality in the time course of the semi-diphthongs /e/ and /o/.

Naeimeh Afshar & Vincent J. van Heuven:
Perceptual Assimilation of American English Vowels
by Monolingual and Bilingual Learners in Iran
Argumentum 18 (2022), 172–191
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/9

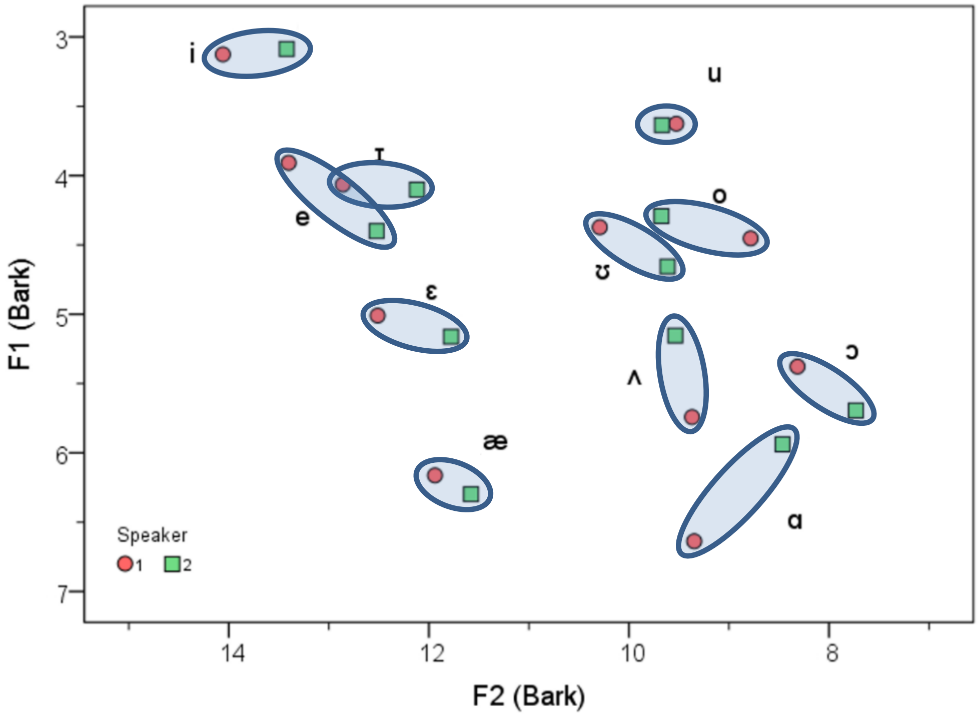

Figure A1. Vowel tokens of Table 1 plotted in the acoustic vowel space defined by F1 (top to bottom, Barks) and F2 (right to left, Barks). Ellipses were drawn by hand and have no theoretical status.

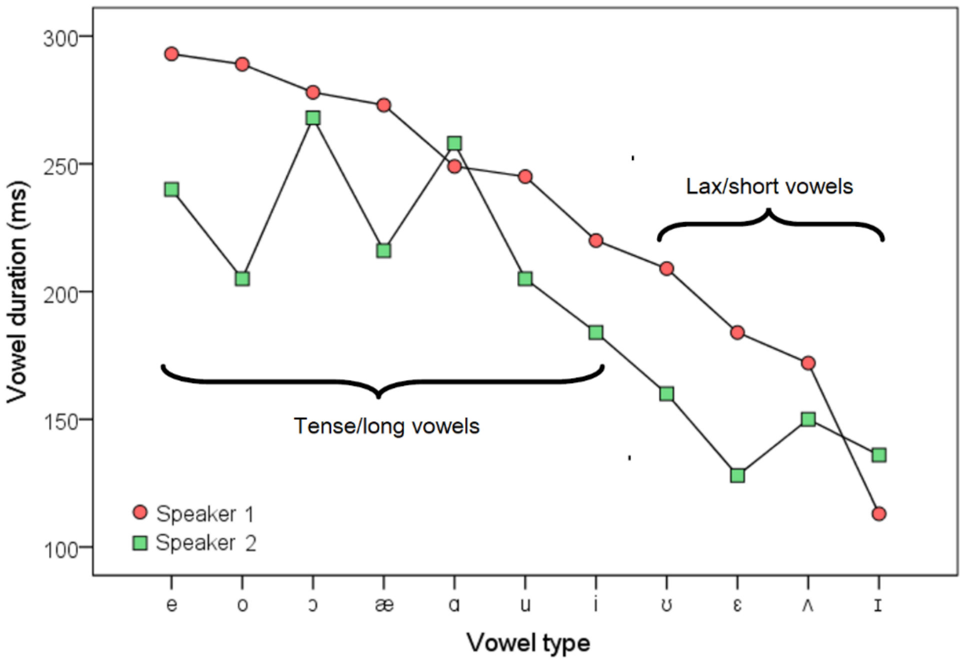

Figure A2. Duration (ms) of 11 American English monophthongs produced by two male native speakers. The vowel types are plotted from left to right in descending order of the duration realized by speaker 1.