The Perception of English Vowels by Native Farsi Speakers

Mehran Mirzazad, 285214

Final Thesis

English Language and Culture

Spring 2021

ITÄ-SUOMEN YLIOPISTO – UNIVERSITY OF EASTERN FINLAND

Philosophical Faculty	**UEF**			
MEHRAN MIRZAZAD				
The Perception of English Vowels by Native Farsi Speakers				
Main Subject	**Työn laji – Level**		**Spring 2021**	**66**
English Language and Culture	Pro gradu -tutkielma	*		
	Sivuainetutkielma			
	Kandidaatin tutkielma			
	Aineopintojen tutkielma			

Abstract

Second language learning has been the subject of many studies in recent years but there have not been sufficient studies regarding those Iranians who strive to learn the English language. With regards to the English language in Iran, there is no L2-speaking community and English is taught and learnt in a non-naturalistic setting.

The purpose of this study is to find out the extent to which Perceptual Assimilation Model (PAM) and Second Language Linguistic Perception would predict the difficulty of learning Standard Southern British English (SSBE) vowels by Persian (FA) speakers. To achieve this, the difficulty of learning English vowels for naïve FA listeners of English were measured by studying cross language categorization of L2 to L1 vowels categories, made possible by perceptual assimilation of L2 to L1 vowels. Following on from this, it was tested to identify whether PAM could predict the level of difficulty of L2 contrast discrimination since PAM regards vowels in pairs.

Then, the results were reported. As it turns out, perceptual assimilation results were partly consistent with those of acoustic predictions. This study shows that perceptual assimilation model has some strength in predicting difficulties when it comes to learning L2 vowels.

Keywords

PAM-L2, L2LP, Linear Discriminant Analysis

Table of Contents

1 Introduction

My findings on second language acquisition have proven to me that experience with another language has more or less some influence on perception of vowels in a target/second language. This study aims to estimate if and to what extent this will pose a problem to Iranian/Farsi learners of English as their second language (L2). Usually, adults who try to learn a second language (L2) have difficulty in dealing with the L2 sounds; that is, acquiring L2 sounds is problematic for the learners. It is interesting to know infants who have innate ability to discriminate sounds of any language lose this ability in 6 months to a considerable degree (Colantoni, Steele & Escudero, 2015). Hence there are a few models that could explain why perceiving L2 sounds are difficult for learner(s) such as the speech learning model (SLM: Flege 1995). Another model which is important for this study and would constitute the framework within which this study has been done is the Perceptual Assimilation Model (PAM) for second language learners developed by Best (1994a, 1994b) which was further developed as PAM-L2.

This model takes into account the notion of difficulty level by seeing the interaction between language learners' first language (L1) and second language (L2) sound systems (Best & Tyler 2007). To evaluate this, there should be an assimilation test/task where participants will try to categorize L2 sounds into their own language category (L1 category). Some categorization pattern will be obtained through this task which will pave the way for predicting the level of difficulty in perceiving second language sounds. Therefore, the aim of this study is to look into PAM-L2's predictions in studying the difficulty of Iranian speakers of English whose L1 is Farsi and L2 is English. To achieve this goal, this study aimed to support the results of an assimilation task with the results of a categorical discrimination task before comparing the two with the results of acoustic task following Ghaffarvand Mokari & Werner (2018). The first and second task have been carried on native Iranians who had little knowledge of English language, therefore could be called naïve learners of English.

Best's Perceptual Assimilation Model (PAM) was originally proposed to predict the difficulty of naïve listeners within a framework (Best, 1995). The idea is that not all L2 categories are equally difficult or easy for naïve learners (listeners). It suggests this difficulty can be examined in terms of how an L2 phonemic contrast is assimilated to L1 categories.

It is supposed that adult listeners have difficulty in discriminating phonetic distinctions which do not sound like their native phonological contrasts. An example of such perceptual difficulty is the poor discrimination of English /r/-/l/ by speakers of Japanese (Goyo, 1971).

A few studies have proven that the inventory size of L1 vowels plays a role when discriminating L2 vowel. In other words, if L1 has bigger vowel inventory size compared to L2, it is highly likely that learners will have an easier time discriminating L2 vowel contrasts. Therefore, vowel inventory size affects the establishment of these categories in learners' minds.

When thinking about acoustic aspects of L2 you have to take formants into consideration. Discerning a vowel from another is made possible by the difference in formants. Also taking formant values into consideration is important when we talk about vowels in a language. There can be two reasons for this (Hunter & Kebede, 2012): First, for identifying vowel in a correct way (Intellegibility). Second, for the naturalness of speech or native-like quality of the vowel. Although different languages may employ different vowels and may consequently have different acoustic properties (including formant values), first (F1) and second (F2) formant values are the important ones.

There are many studies in which PAM-L2, L2LP and SLM have been used, however there has been no similar study on Farsi speakers' difficulty in learning English. On the other hand, there has been studies regarding Azerbaijani speakers of English (Ghaffarvand & Werner, 2018) as a large demographic population in Iran.

1.1 Aims and Scope

This study has aimed to predict difficulties of Iranian learners of English in production and perception of L2 vowels. The predictions have been done by performing a Perceptual Assimilation task based on PAM-L2 and L2LP and in comparison with acoustic data from another study by Ghaffarvand Mokari & Werner (2017). Result of assimilation task has been supported by having a forced-choice discrimination task (task 2). The second task consists of calculation of sensitivity A' score for each contrast/vowel pair. For the acoustic part of this study, a linear discriminant analysis has been used for the purpose of finding out degree of correct classification in distinction of 6 Farsi vowels and 11 Standard Southern British English (SSBE) vowels.

Having mentioned earlier, the size of first language (L1) vowel inventory plays a role in the perception of second language (L2) vowels. This means, if L1 has bigger vowel inventory than

L2 it can presumably be easier for learners to perceive and discriminate L2 vowels. In the same way, one can imagine if L1 vowel inventory size is smaller than L2 vowel inventory, discrimination can be challenging for learners. Based on this, we have meant to explore answers for the research questions mentioned above.

FA (Farsi) has six monophthongs including /æ/, /o/, /e/, /u/, /i/ (Ghaffarvand Mokari and Werner, 2017) and SSBE on the other hand has eleven monophthongs including /æ/, /ɑ/, /ɒ/, / ɛ /, /ʌ/, /ɜ/, /ɔ/, /ʊ/, /u/, /ɪ/ and /i/ (Roach, 2004). Below are the FA and SSBE vowels presented in the traditional cardinal vowel charts. As it can be seen in Figure 1, FA has a much smaller vowel inventory than SSBE where many of the SSBE vowels do not have a counterpart in FA. Regarding the substantially different vowel systems of the SSBE and FA, it can be expected that discrimination of certain L2 contrasts would be challenging/difficult for FA listeners.

Figure 1. Vowel Charts of SSBE (left) & FA (right) based on IPA illustrations in Journal of International Phonetic Associations.

The acoustic representations of FA and SSBE vowels are shown in a Bark-scale in Figure 2. The acoustic representations are based on productions of ten male speakers from the study by Hawkins and Midgley (2005) for SSBE vowels and the productions of 28 male speakers form the study by Ghaffarvand Mokari and Werner (2017) for FA vowels recorded in /hVd/ contexts.

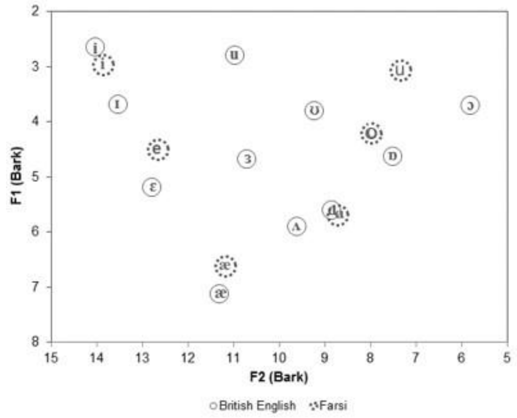


Figure 2. F1-by-F2 plots FA and SSBE based on male speakers' productions from study of Hawkins and Midgely (2005) for SSBE vowels and Ghaffarvand Mokari and Werner (2017) for FA vowels.

1.2 Research Questions

For this study, a null hypothesis and four research questions have been assumed:

Null Hypothesis: Native Farsi speakers are not able to distinguish vowels which are not part of their L1 vowel inventory at all.

Research Question 1: Which English vowel pairs will be difficult for Farsi speakers?

Research Question 2: How accurate Farsi speakers can discriminate English vowels based on perceptual similarities of L1 and L2 sounds?

Research Question 3: How will Farsi speakers categorize English vowels using their mental representation of L2 sounds?

Research Question 4: How similar Farsi and English vowels are from an acoustical standpoint? Do cross-linguistic acoustic similarities match perceptual assimilation or not?

2 Literature review

2.1 Farsi and English vowel systems

Models such as SLM and PAM and L2LP suggest that understanding of non-native vowels can be rather determined by extent to which these new sounds can be related to speaker's first language (L1) (Bion et al. 2006).

Vowel Structures can be complex (Ladefoged & Johnson, 2011). We may think of vowels in terms of containing different pitches at the same time. There are two types of pitch: a) The pitch at which vowels are spoken which depends on pulse being produced by vibrating vocal folds, b) Overtone pitches that are dependent on resonating cavities of vocal tract's shape. Of these two, overtone pitches are the ones that give a vowel its distinctive quality. Usually separate overtones of a vowel cannot be heard as distinguishable pitch, however there are certain circumstances which will make it possible for the overtone pitches to be heard.

Taking the following words as an example *heed, hid, head, had, hod, hawed, hood, who'd* and trying to whisper them will make it clear that sounds of these set of words are formed on a descending pitch. These characteristic overtones are the highest for the vowel in *heed* and the lowest for the vowel in either *hawed*, *hood* or *who'd*. The Pronunciation for *heed, hid, head, had, hod, hawed, hood, who'd* in British accent have been provided by Ladefoged & Johnson (2011). For the first four of these words, vowels quality goes up in pitch while the last four words have a declining pitch. From the articulation perspective it refers to *vowel height.* The aforementioned characteristics are called ***formants***. The first formant (F1) refers to the one with the lower pitch and the second formant (F2) refers to the one with the higher pitch.

Farsi (Persian) is the official language of Iran with population of over 80 million people. In this study Persian/Farsi vowels have been addressed as *FA*. FA has already been used as a representative for Iranian learners/speakers of English. Persian language is also spoken in Afghanistan and Tajikistan with Dari and Tajiki dialects respectively. Persian is looked at as one of Indo-Iranian subgroups of the Indo-European language family (Hudson, 2000). In contrast, English is one of Germanic subgroups of Indo-European language family. The followings are the formant values for SSBE and FA:

Table 1. Mean Formant frequencies of the eleven phonetic monophthong vowels of SSBE as produced by three L1 speakers (Hunter & Kebede, 2012)

	Formants	
SSBE Vowel	F1 / Hz	F2 / Hz
iː	296	2241
ɪ	396	1839
e	532	1656
æ	667	1565
ʌ	661	1296
ɑː	680	1193
ɒ	643	1019
ɔː	480	857
ʊ	395	1408
uː	386	1587
ɜː	519	1408

Table 2. Mean Formant frequencies of the six phonetic vowels of Farsi as produced by 12 young female L1 Farsi speakers (Hunter & Kebede,2012)

	Formants	
Farsi Vowel	F1 / Hz	F2 / Hz
i	365	2508
ɛ	644	2115
æ	990	1722
ɒ	750	1251
o	558	1102
u	423	1065

2.2 Previous approaches to vowel assimilation

In order to have a reliable measure of cross-linguistic similarity we need to perform a variety of tasks (Cebrian, 2006); some of which are vowel assimilation tasks and some are discrimination tasks upon which much of L2 perception research and studies are based (Escudero & Chladkova, 2010). In perception studies, cross-linguistic assimilations tasks are the ones that lay the foundation for these studies (Cebrain, 2008; Rallo Fabra, 2005; Rallo Fabra & Romero, 2012). Therefore, when classifying L2 sounds in terms of L1 sounds, using cross-linguistic assimilation tasks will be inevitable. To support the findings, participants may be asked to judge their perceived goodness by assigning a goodness rating usually on a numerical scale after classifying the sounds (as is done). This enables the researchers to make predictions on relative difficulty in discrimination of two L2 sounds within the framework of PAM-L2.

2.3 Second language learning models

The term cross-language speech perception refers to those models which explain the difficulties of learners in learning a second language (L2). Different models have been suggested for this issue such as Speech Learning Model (SLM) (Flege, 1995) and Perceptual Assimilation Model (PAM) (Best, 1995). PAM, as being mentioned before, accounts for perception and discrimination of non-native contrasts by naïve learners. Hence it is necessary to briefly explain SLM, PAM-L2 and L2LP models to let reader(s) gain a general understanding of each model and principles of each.

2.3.1 Speech learning model (SLM)

Speech learning model (Flege,1995, 2003) is concerned with L1 and L2 plus relationship between L1 and L2. Key aspects of SLM are taking into consideration the level of success or failure of learners in both perception and production in regards to vowels and consonants. In this model perception of L2 sounds precedes production of those sounds. Within this framework, the ability to understand phonetic difference of L1 and L2 also depends on factors such as to what degree one has experience with an L2 and at what age learning of the L2 has appeared to begin. This model makes use of individual sound segments compared to PAM which takes L2 vowel pairs into consideration where emphasis is being put on perception of those vowels. Flege (1995) states that if learners of L2 have received enough input, they may perceive L2 phonetic properties very well. Moreover, nature of input is important. It is suggested by Flege that much like the development of L1, production of L2 is led by those perceptual representations that are stored in learners' long-term memory. SLM maintains that the ability of acquisition of L1 achieved by mechanisms such as the ability to form new phonetic categories, will remain accessible and intact through one's lifetime. This model assumes a phonological space in which phonetic elements of L1 and L2 subsystems have influence on one another. SLM hypothesizes that the less similar an L2 sound to an L1 sound, it is more likely for it to become a new category by learners. In the same manner, when an L2 sound is very similar to an L1 sound where no new category formation happens, there will be assimilation of L1 and L2 sounds creating a *merged* L1-L2. Furthermore, SLM posits that when the size of vowel inventory in L1 is smaller than L2, ability of learners to distinguish differences of L1 and L2 will increase. Since our current study has been done within the framework of PAM-L2 and L2LP, our results will not cover SLM.

A study by Flege, Takagi & Mann (1996) investigated to see how would Japanese learners (L1) fare in perceiving /ɹ/and/l/. To be more accurate, 36 Japanese participants were tested. Their lexical familiarity and experience with English language were taken into consideration as possible factors affecting identification of word-initial tokens of English /ɹ/ and /l/. The inexperienced Japanese participants shown as (IJ) had lived in the U.S. for 2 years, while the experienced ones shown as (EJ) had lived there for 21 years on average. 42 words, such as *road, load*, etc. and four non-words such as *rine, ruck*, etc. were presented to learners. For most part, experienced learners did better than inexperienced learners. They were asked to distinguish /ɹ/ and /l/ when the words were presented to them. The context in which consonants were presented to the learners had some effects on their performance. Where /ɹ/ and /l/ were presented to the learners in isolation, despite lack of lexical information, the experienced Japanese learners performed well. Based on SLM, native-like performance achieved by learners would mean the existence of new L2 categories for /ɹ/ and /l/ which makes category formation possible to learners. The results also show that when realisation of difference between L1 and L2 takes longer time as a result of L1 and L2 influence, it would mean that L1 and L2 are stored as a single category also known as *Merged-Category*.

2.3.2 Perceptual assimilation model (PAM)

This model rules out storage of categories as mental representation and instead it puts forth a direct realist approach in which articulatory gestures are deemed to be involved regarding detection of information by listeners. Compared to SLM, the predictions within this framework are made regarding contrasts/vowel pairs rather than individual L2 phonemes (Best 1995). This model has been originally developed to address learning of target language non-native contrasts by listeners who are naïve and with no prior experience of TL language. It was further developed by Best and Tyler (2007) to incorporate studies on perception of L2 by leaners (PAM-L2). What this model has in common with the SLM model is that both address possibility of unequal level of difficulty in discrimination of a TL contrast. Therefore, PAM approaches learning of L2 with prediction for discrimination of contrasts in a way that the learner's ability in discerning TL/L2 contrasts will be depending on how each member of a contrast is assimilated. This will lead to different category formations: Single-Category (SC), Two-Category (TC), Category-Goodness (CG), Uncategorized-Uncategorized (UU), Categorized-Uncategorized (CU) and Non-Assimilable (NA). Best and Tyler (2007) regard TC, SC, CG and CU as four distinct assimilation patterns in this framework. In most studies carried out within the framework of PAM-L2, the emphasis has been on SC, TW and CG assimilation types.

To put PAM into practice, 22 native speakers of American English (AE) were tested for discrimination of Zulu and Tirinya as target languages (TL) by Best, McRoberts & Goodell (2001). This study demonstrated that contrasts falling in Two-Category assimilation patterns will be easier to discriminate compared to other assimilation patterns in spite of absence of both members of a contrast in L1. This study included two tasks: one assimilation task and one discrimination task where participants had to decide whether the sound they hear (X) is the same as A or B.

2.3.3 Second Language Linguistic Perception (L2LP)

This model (Escudero & Boersma, 2004; Escudero, 2005, 2009) has been proposed to account for difficulty of learners of various ages being at different levels of proficiency in learning a TL or L2. Non-native learners can be people who are highly trained and experienced in using TL/L2 or people who have just begun learning a TL/L2. Within the framework of L2LP, categories of learners' L1 are compared with those of the language they aim to learn. It has been made clear that although this model can account for different learners who are at various stages of learning a TL/L2 language, naïve learners are favoured to be investigated. One thing that both PAM and L2LP have in common is taking into account vowel contrasts rather than individual segments. L2LP accounts for acoustic aspects of TL/L2 in addition to perceptual similarities in relation to difficulties of learning L2. Elvin et. Al (2014) mentions prediction for L2 learning difficulty by taking into account the acoustic properties of L1 and L2. Based on

Colantoni et. al (2015:44), it will be a better option to have longitudinal data gathered from naïve learners/listeners whose acquisition of L2 gradually takes place over time although L2LP can incorporate learners who are at various levels of proficiency.

Regarding different scenarios possible within this framework, three scenarios have been suggested by Escudero (2005): *NEW, SUBSET and SIMILAR.*

A NEW scenario refers to more categories in L2 than in L1. In L2LP when a contrast falls in this category it will be the most difficult contrast to be perceived by L2/TL learners. Falling in this category means an L2 (non-native) contrast has been assimilated to only one L1 category.

A *SIMILAR* scenario suggests existence of the same categories in both L1 and L2. A SIMILAR scenario is easier than a SUBSET scenario. In this scenario, there is an assimilation of an L2 contrast to two different L1 categories. L2LP suggests a contrast falling in this category will be the easiest to perceive.

A SUBSET scenario marks existence of more categories in L1 than L2. A SUBSET scenario also known as multiple category assimilation is easier than NEW scenario since a contrast falling in this category will be perceived by L2/TL learners with less difficulty. In this scenario a non-native L2 sound is perceived as more than one category in L1.

In a study by Escudero & Boersma (2004), they investigated perception of Scottish English and Standard English by Spanish learners. Their study has been narrowed down to two vowels: /i-ɪ/ in relation to different acoustic values of the two vowels. The study has found Spanish learners of the two varieties perceive and keep acoustic values of these two varieties according to their L1. Perception of learners of SE were native-like due to resemblance of formant frequencies of Spanish /i/ and /e/ to the two varieties of English. On the other hand, learners of SBE have a tendency to categorize vowels based on their duration differences. Their study has concluded that acoustic properties of TL/L2 does actually have an influence on learner's performance in perceiving vowels.

3 Previous studies

I found no previous study conducted solely on examining Farsi speakers' difficulty in learning English vowels by the time we started this project. As a result, the centre point of this study has been placed on analysing and determining the scope of difficulty FA speakers experience in learning English vowels within the framework of PAM-L2 and L2LP. The following studies will be discussed in the section designated to previous studies (Theoretical Background):

(1) Gilichinskaya & Strange (2010): Perceptual assimilation of American English vowels by inexperienced Russian listeners

(2) Chladkova and Podlipsky (2011): Native dialect matters perceptual assimilation of Dutch vowels by Czech listeners

(3) Angelos Lengeris (2009): Perception of Standard Southern British English vowels by native speakers of Japanese and Greek language

(4) Andrew Jeske (2007): Perception of English vowels by native Spanish speakers

(5) Winfred Strange et al. (2004): Perceptual similarity of North German and American English

3.1 Perceptual assimilation of American English vowels by inexperienced Russian listeners

In a study by Gilichinskaya & Strange (2010), the perceptual assimilation of American English vowels by inexperienced Russian listeners was studied. Participants' L1 was Russian and their L2 was English. They selected these participants because they were monolingual speakers of the Russian language. This was for the purpose of minimizing effect of L1 phonology on perception of L2 sounds. The aim of their study was to: (a) find out if acoustic similarity of Russian and American vowels can predict their perceptual assimilation to Russian vowels by those monolingual Russian speakers and how accurate it can be. (b) find out the most difficult contrasts that are perceived within the framework of PAM.

For their method, they chose a male native speaker of American English to produce these 8 vowels in their study: /i:/, /ɪ/, /ɛ/, /æ:/, /ɑ:/, /ʌ/, /ʊ/, /u:/. Their participants were 19 native Russians between 19 to 68 years of age and they all signed informed consent forms. They all reported normal hearing capability except one who let the researchers know about it on the informed consent form and consequently was excluded from the data analysis. Some of their participants were researchers whilst others were employees and also graduates. Two of their participants had been previously exposed to other languages. They were tested in a quiet room. The software used for the test was called Paradigm 4.0 with other devices included headphones and laptop. Participants had to listen to meaningless English words with the emphasis on the first vowel to see which Russian vowel or combination of vowels sound similar to the stimuli. The responses that had to be chosen by participants were written in standard Russian orthography. After selecting responses by the participants, the perceived goodness rating was measured on a scale from 1 to 9.

The results were presented in form of a table. The American English and Russian vowels and the percentage for frequency of responses (the Russian vowel category that was selected) were shared. The number of times each participant selected a certain response was presented as well. Also, the goodness of ratings perceived by individuals were shown on a table in the result section. In general, their results have concluded that American English /i:/, /ɑ:/, /u:/ were consistently assimilated by participants to Russian /i/, /a/ and /o/ respectively. At a moderate level of consistency, American /ɪ/ and /ʊ/ were assimilated to Russian /ɛ/, /o/. The least consistently assimilated American vowels were /ɛ/, /æ:/, /ʌ/ that were assimilated to Russian /ɛ/, /o/.

They compared these findings to acoustic similarities between American English and Russian vowels and a linear discriminant analysis was performed. They used a classification matrix of first, second and third formant bark values (F1, F2, F3) based on data gathered in a previous study by Gilichinskaya et. al (2007). Their results pointed out that American English /i:/, /u:/ were the most similar to their Russian counterparts acoustically, while American English /ɛ/, /æ:/, /ɑ:/ and /ʌ/ were most similar to a single Russian /ɛ/ vowel. The acoustic results were in line with the results of perceptual assimilation task since the patterns for three more experienced Russian listeners were no different than those of inexperienced Russian listeners.

Based on these patterns, they concluded that there is no difference between those participants who were more experienced with the English language. As a result, their exposure to English has been deemed as being irrelevant to this study by researchers. According to their findings, the most difficult contrasts perceived by Russian learners of English within the framework of PAM were AE /ɪ-ɛ/ and /ɑ:-ʌ/ categorized as Single Category, AE /æ:-ɑ:/ as Categorized-Uncategorized and AE /i:- ɪ/ and /ʊ-u:/ categorized as Two-category which will be quite easy for Russian learners to perceive.

3.2 Native dialect matters perceptual assimilation of Dutch vowels by Czech listeners

In another study by Chladkova and Podlipsky (2011) titled N*ative dialect matters perceptual assimilation of Dutch vowels by Czech listeners*, the importance of certain dialects on perception of second language learning was demonstrated. It was discussed that prior experience with L1 will limit a learner's capacity to perceive and produce L2 sounds. Reference to Flege (1995) was given who discussed SLM model with regards to similarity and dissimilarity of L2 sounds to closest L1 counterparts. There was also reference given to Best (1995) and Best & Tyler (2007), based on whom PAM and PAM-L2 can predict the degree of difficulty for L2 sounds.

Chladkova and Podlopsky (2011) mentioned the similarity of L2LP and PAM in labelling assimilation categories. For example, both PAM and L2LP predict a single-category assimilation for a case in which L1 has smaller inventory compared to inventory size of L2. As for different dialects affecting the perception of L2, they have referred to Escudero and Boersma (2004) and Baker and Smith (2010). Escudero and Boersma's results shows more difficulty in perception of /i:-ɪ/ contrast for Spanish learners of Scottish English in comparison with difficulty of Spanish learners of Southern British English due to the fact that /i:-ɪ/ are not contrasted by duration in Scottish English. Baker and Smith demonstrated the difference in perception of French /i/, /y/, and /u/ between English learners of Quebec French and English learners of European French.

In this study, Chladkova and Podlipsky examined whether or not the speakers of the same language with two different dialects and prior experience with the other dialect would show difficulty in the perception of L2 sounds. The language was Dutch and two groups of participants were selected. Participants had different dialects of Czech language: Bohemian Czech and Moravian Czech.

With regards to method, they had 41 Czech speakers, 19 of whom were users of Bohemian dialect with the rest being users of Moravian dialect. The Bohemian group consisted of 5 males and 14 females aged 20 to 33. The Moravian group consisted of 7 males and 15 females aged between 20 to 35. The mean of their ages has been given as well. The participants were all university students who had never lived in another country for more than one month with no prior exposure to the Dutch language while they all had experience with the other dialect. In other words, the participants of the Bohemian group had frequent exposure to Moravian language while the other group had frequent exposure to Bohemian dialect by different means.

The stimuli used in their study were recorded by a native male and female aged 23 and 22 years respectively. Their stimuli were in /hVb/ context which is possible in both Czech and Dutch phonotactically based on the researchers. They performed a multiple forced-choice identification task involving 96 monosyllables that were presented randomly. Their participants were tested in a quiet room using headphones. On the computer screen they showed their participants were exposed to ten words while the stimuli were being played which were nonsense words that could also possibly be Czech nouns spoken by male and female native speakers. The participants had also been offered a round of practice with the aim of giving them an opportunity to familiarize themselves with the test and the nature of stimuli.

In their result section, a table was provided to present the perception of Dutch vowels by naïve Czech listeners divided by the dialects which shows similarity between Bohemian and Moravian choices. Mean percentage for assimilation of Dutch vowels to L1 categories has also been given. To determine whether there were differences between Bohemian and Moravian responses to stimuli, they carried out t-tests.

Based on Chladkova and Podlipsky, Moravian listeners and Bohemian listeners were different in perceiving Dutch high round vowels as they weigh durational and spectral cues to the Czech /iː-ɪ/ contrast. They discussed that these two groups of participants were different when it came to perceiving Dutch high-front vowel explaining that Dutch /i/ was assimilated to Czech /ɪ/ by Moravian listeners while a short Dutch /i/ was perceived as Czech long /iː/ by Bohemian listeners. This study concluded that Bohemian and Moravian listeners had different perception of Dutch /y/ which is a high front rounded vowel. As presented on the table, this vowel was perceived as /iː/, /ɪ/, /uː/ and /u/ by the participants. Their results confirmed there were a few single-category assimilation patterns. For example, on more than 50 % of times, Dutch mid-high front /ɪ/ was perceived as Czech short high-front /ɪ/. Their study successfully demonstrated that Bohemian listeners had more difficulty compared to Moravian listeners in perceiving Dutch /y-ʏ/. These findings were in line with L2LP model suggested by Escudero (2005) which explains that difference in listener's native dialect can affect perceptual assimilation of some L2 vowel contrasts. In addition, from an acoustic perspective the between-dialect differences in L1 resulted in differences in the perception of L2 sounds.

3.3 Perception of Standard Southern British English vowels by native speakers of Japanese and Greek language

Lengeris (2009) investigated perception of Standard Southern British English vowels by native speakers of Japanese language and Greek language. The results show that both Japanese and

Greek speakers were sensitive to durational cues while perceiving English vowels. Lengeris mentioned “they aimed to examine use of durational cues in the perception of English vowels by the two Japanese and Greek groups as the two languages have relatively similar vowel qualities”. Although there is outstanding difference concerning the use of duration in contrasting vowels.

The difficulty of Japanese learners in perceiving L2 sounds in learning English was mentioned for the purpose of referring to a study by Goto (1971). Japanese learners of English had difficulty perceiving English /r/-/l/ contrast as they rely on second formant (F2) when trying to distinguish /r/ from /l/ (Iverson et al, 2003). There is also mentioning of a hypothesis called feature hypothesis by McAllister (2002) that revolves around the use of duration in L1. Based on this, usage of durational cues in dealing with L2 pertains to usage of duration in L1. The researcher also talked about the problem of Korean learners in L2 learning due to the fact that they lack experience in using duration in their L1 distinction (Flege et al., 1997). Therefore, this study aimed to further explore use of duration as a perceptual cue in L2 learning.

Greek has five vowels /i:, e, a, o, u/ while Japanese has /i, e, a, o, u/ as tense (long) vowels and /i:, e:, a:, o:, u:/ as lax (short) vowels. To achieve their goal, two tests were conducted: a perceptual assimilation task and a categorical discrimination task. This would support identifying whether cross-language assimilation would predict discrimination of English vowels.

With regards to the method, three female speakers of Southern British English, all native, were asked to record the English stimuli. The mean and range of their age were given by the authors. They produced 11 English monophthongs in the sentence *I read ___ on the screen* and the vowels were produced in two /bVB/ and /bVp/ context. Three participants read the 11 vowels on two contexts four times totalling 264 tokens. The brands for the recorder and the microphone were stated. This process took place in an anechoic room at the University College London. The first and second formants (F1, F2) were measured using SFS speech analysis.

For the recording of Greek and Japanese vowels, there were three female native speakers of the Japanese language and 3 female native speakers of the Greek language. They produced their L1 vowels in a /mVn/ context uttering the sentence *I read __ on the screen.* Once again the mean and range of their age was given. They produced 60 tokens for the Greek language, that is 3 speakers used 5 vowels in 4 repetitions and 120 tokens for the Japanese language which means 3 speakers used 10 vowels in 4 repetitions. The recording was done with a recorder (the name of the recorder was given) with the process taking place in a quiet room. The first three repetitions of each vowel were chosen to be used for acoustic analysis.

For cross-language perceptual assimilation task, they had 33 participants, of whom 18 were native Greek speakers and 15 were native Japanese speakers. The mean and range of their age was given and their localities in their respective countries was also mentioned. They were tested in London during their stay for a summer course and had 10 to 15 years of experience with L2 (English language) through studying with L1-accented teachers. Their level of English language was rather high (having good grades in IELTS or TOEFL) however, all participants had little or no interaction with native speakers of English. None had lived more than a month in an English speaking country according to the questionnaires that were completed before the actual test took place. There were no reports of any hearing deficiency or language impairment cases.

Their results for the assimilation task showed that Greek listeners assimilated more than one English vowel to the same L1 category while Japanese listeners assimilated each English vowel to a different L1 category except English /ɜ:/ and /a:/ which were both assimilated to Japanese /a:/.

They conducted a discrimination task to: (a) determine if cross-language perceptual assimilation patterns could predict L2 discrimination in the framework of PAM (b) to examine discrimination of English vowels in /bVb/ and /bVp/ context by both Greek and Japanese listeners. They basically used the same participants who were all born in London. Based on previous studies on Spanish learners of English, some contrasting vowel pairs were selected to be used for this discrimination task since Spanish vowel system is similar to Greek vowel system.

Every contrast was tested through 8 trial times containing an odd vowel category and 8 catch trials as well, which included three tokens of the same vowel category. The objective was for the participants to identify the odd one out and click on the number 1,2 or 3 in different trails and *same* if judged to be from the same category. Lengeris mentioned “the sensitivity A' score (Snodgrass et al. 1985) was calculated for discrimination of each vowel pair based on hits when the odd one was selected in different trials and false alarms when it was incorrectly selected in catch trials”. It was done as follows:

If H (hit) = FA (false alarm) then A' = 0.5

If $H > FA$ then

$A' = 0.5 + [H - FA) * (1 + H - FA)]/[(4 * H) * (1 - FA)]$ and

if $H < FA$ then

$A' = 0.5 - [(FA - H) * (1 + FA - H)]/[(4 * FA) * (1 - H)]$.

The chance level would be indicated by A' score of 0.5 while the perfect discrimination of a contrast would be indicated by A' score of 1.0.

Two-way analysis of variance (ANOVA) yielded considerable effect of Contrast and also considerable effect of Context plus Contrast × Context which proved the sensitivity of Greek listeners to vowel duration. For Japanese participants, the ANOVA also yielded considerable effect of Contrast plus considerable Contrast × Context interaction. Lengeris mentioned "for both groups of participants, the effect of context on discrimination scores was not uniform across contrasts". The results of the second task were in line with the results of the first task indicating that both groups use durational cues when trying to discriminate L2 vowels.

The similarity of L2 groups' vowel systems, vowel structure and phonotactic constraints were the reason why the researcher selected these two groups. The results of their first task showed that except /ɜː/ and /aː/ which were related to different L1 spectral category, most English vowels were assimilated to the same spectral categories.

In short, this study aimed to investigate if effectiveness of prior experience with duration in L1 vowels can be related to the perception of vowels in L2 regarding the use of durational cues. The languages under study were Greek and Japanese because both have similar and simple vowel structures. In terms of the stimuli, vowels in this tasks were uttered in two contexts: /bVp/ and /bVb/. The reason for choosing these two languages was the similarity of their vowel inventories, simplicity of their syllable structures and phonotactic constraints.

The first task provided a positive answer to the research question of whether durational cues are available to learners who have no such L1 experience. However, this task proved that English vowels will better fit into the L1 categories when they are in a context that resemble the duration of the L1 vowel which is spectrally the closest to it.

The second experiment was in accordance with PAM's framework. The Single-Category contrasts were the most difficult one to be discriminated by the learners who had less difficulty in discriminating Uncategorized-Categorized contrasts and also less difficulty with Category

Goodness contrasts. On the other hand, learners discriminating Two-Category contrasts had done so with ease and no difficulty.

Their result showed both groups were sensitive to durational cues. When the durational difference across contexts was the same, some contexts were discriminated easier in one context than the other. Although, there were cases of context-induced changes in duration difference between two vowels which did not lead to a change in discrimination.

Lengeris mentioned “temporal cues in L2 may be available to L2 learners who do not use duration if the L1 duration category does not interfere with the perception of an L2 contrast temporally”. They concluded that regardless of which contrast is to be perceived, lack of temporal difference for the English /iː-ɪ/ contrast might be the reason for this general view that durational cues are available to learners with single vowel duration category. They concluded the results of their study are compatible with PAM-L2. Researchers explained that vowel perception is made possible through a sophisticated interaction between relationships of L1 and L2. Their final conclusion summarized that ways of accessing spectral and temporal information in L2 seem to be more or less the same across people who have different L1 backgrounds.

3.4 Perception of English vowels by native Spanish speakers

Andrew Jeske’s thesis was on the perception of English vowels by native Spanish speakers. Spanish has five vowels: /ɪ, e, a, o, u/.

For a comparison between the two languages, he presented the formant values of English and Spanish on a table which was from three previous studies completed by other researchers. The aim of his study was to add to pre-existing knowledge on the perception of English language by Spanish learners within the framework of PAM-L2. In doing so, he led three tasks. These 3 tasks within the framework of PAM-L2 were to determine the difficulty of Spanish learners of English. In other words, how Spanish learners perceive English vowels and which vowels will be more difficult to perceive.

With regards to method, for the recording of stimuli for vowel assimilation task, four native speakers of Standard American English were selected. Two of them had clear voices and the recordings of their voices were thus favoured over the other two. The distinction of /ɔ/ and /o/ was made possible due to the clarity of the voices of these two Speakers. The words were recorded by a USB microphone in Media Centre of University of Pittsburgh.

Before the actual test began, the participants were given instructions on the screen and sufficient time given for studying it. They also had the opportunity of 5 trial sessions to be better familiarised with the actual test. After the practise session, they would move on to the actual test or spend even more time on the training session to be more acquainted with the test. The stimuli were randomly presented while using headphones. When an English word was presented, the participants had to match the vowel to one Spanish vowel by choosing the correct response on the screen. This involved several Spanish words that contained the vowels. When they did classify the vowel, they heard it again and this time they had to select on a Likert scale of 1 (for not similar) to 9 (for very similar) the degree of similarity of the vowel they had just heard.

Their participants were 5 males and 7 females, all native speakers of Spanish. Their nationalities were asked and the range of their age was given. For the purpose of clarification of their experience with the English language, they were asked to mention the number of years they had studied English. They were all university graduates who had normal hearing and English was their most fluent L2.

As for the result of the first task, the researcher presented a table that contained the mean percentage of times that each English vowel (ten vowels in total were under investigation) was assimilated to each Spanish vowel. The goodness of ratings was also presented however, percentages less than 4% were not shown.

The vowels with the highest percentage of assimilation were /i/ and /æ/ with 93% assimilation to Spanish /i/ and /a/ respectively. English /i/ had 93% assimilation and English /ɪ/ had 75% assimilation to Spanish /i/. This proves the distinction of /i-ɪ/ is difficult for the participant as it is a case of Single-Category assimilation. Another Single-Category contrast was /æ-ɑ/ which proved to be difficult to perceive. There were cases of Category-Goodness e.g. /æ-ɔ/ or /ʊ-u/ for the majority of contrasts thus making them more difficult to perceive compared to Two-Category contrasts and therefore easier to perceive than Single-Category assimilations. There were two cases of Two-Category assimilations e.g. /ɛ- æ/ which proved to be easy to perceive. All these assimilation patterns were presented on a table.

For the second task the researcher used an AX categorical discrimination task to find out which L2 contrasts are difficult to discriminate for the native speakers of Spanish language. The stimuli presented to participants were the ones used in his first task which were in /bVC /and /dVC/ contexts. It was conducted in a Lab at the university of Pittsburgh using headphones.

The participants had to listen to the vowels presented to them and press either a green button or a red button. In short, this task is a *same* or *different* test to determine if the words represented to participants contain the same or a different vowel. Before actual task began, 5 trials were presented to the participants. The twelve participants were the same from the assimilation task. There were also cases where participants had accidentally pressed wrong button. In total, it took 24 times for a vowel to be discriminated.

Exploring the results, among three Single-Category assimilations, /i-ɪ/ was the most challenging since members of this contrast were misidentified by the participants 74% of the time. For /æ-ɑ/ contrast, the participants misidentified the vowels 52% of the time. The easiest notable contrast was /e^{I}-ɛ/ as it was perceived as the *same* for only 13% of the time. For Category-Goodness assimilations, /æ-ɔ/, /o-ʊ/ were easy to perceive since they were perceived as the *same* 39% and 25% respectively. The two /ɑ-ɔ/ and /ʊ-u/ contrasts were perceived as the *same* 61% and 67% of the time respectively. The researcher did not take into consideration /o-ɔ/. For Two-Category assimilations, it was interesting to see /ɛ-æ/ being perceived as the *same* 75% of the time, making this contrast the most difficult one to be discriminated among all other contrasts under investigation. The other contrast of Two-Category assimilation pattern was /ɪ-e^{I}/, which was perceived as the *same* 22% of the time. In total, out of three Single-Category assimilations, one was easily discriminated whilst the other two were difficult. Out of 5 Category-Goodness assimilations, two were poorly discriminated and three were moderately discriminated. Out of 2 Two-Category assimilations, one was easily discriminated and the other one was poorly discriminated. Based on this task, the researcher concluded that irrespective of PAM-L2's predictions any contrast under study could be discriminated well or even poorly. Therefore, the results of the second task revealed that five contrasts were discriminated poorly: /i-ɪ/, /æ-ɑ/, /ɑ-ɔ/, /ʊ-u/ and /ɛ-æ/.

For his final task, Jeske conducted a forced identification task. The method for this task included stimuli from the first task and a practice session to familiarise the participants. They had to choose the right picture on the screen amongst other pictures which corresponded to the word they heard. This was followed by another short practice for the actual test and the first five tokens were not recorded. If the participants wanted an opportunity to repeat the practice test, they had the option to do it or to start the actual test. When participants heard a word containing a vowel such as *bat*, they had to relate it to the picture of a word such as *can* which has the same /æ/ vowel.

The results of forced identification task were then presented on tables in 3 separate categories as the previous tasks.

For *Single-Category* assimilation, once again the /i-ɪ/ contrast proved to be challenging. English /i/ was incorrectly perceived as /i/ only 44% of the time whereas it was perceived as /ɪ/ for 54% of the time. The other member of the contrast was correctly perceived as /ɪ/ 71% of the time.

For *Category-Goodness* assimilation, the /ɑ-ɔ/, /ʊ-u/ contrasts proved to be rather challenging: as /ɔ/ was misidentified as /ɑ/ for 54% and as /ɔ/ for 23% of the time. /ɑ/ was identified as /ɑ/ 63% and was misidentifies as /ɔ/ 18% of the time. /ʊ/ was identified as /ʊ/ 50% and misidentified as /u/ 30%. /u/ was misidentified as /ʊ/ 50% and was identified as /u/ 43%.

For *Two-Category* assimilation, both tokens of /ɪ-e^{I}/ contrast were identified correctly. For /ɛ-æ/ contrast, /ɛ/ was only perceived as /ɛ/ for 57% of the time and was also misidentified as /æ/ for 28% of the time. This was not the case with the other token as /æ/ was only misidentified as /ɛ/ 5% of the time.

To sum it up, the categorical discrimination task revealed that the prediction made via PAM-L2 is not always accurate. Regarding Single-Category assimilations, /e^{I}-ɛ/ was discriminated 87% of the time as opposed to the prediction made by assimilation task. The perceptual assimilation result for this contrast predicted difficulty of perception, whereas in the second task members of this contrast were perceived as the same for only 13% of the time making this contrast easy to discriminate. Regarding Category-Goodness assimilations, /ɑ-ɔ/ and /ʊ-u/ were perceived as the same for 61% and 67% respectively, making them difficult to discriminate whereas /æ- ɔ/ and /o-ʊ/ were easy to discriminate. For Two-Category assimilation, /ɛ-æ/ was perceived as the same with 75% making this contrast difficult to discriminate.

In the third and final task, regarding Single-Category assimilation, /i-ɪ/ was the most problematic contrast introduced to participants, as in this contrast /i/ was perceived as /i/ 44% and as /ɪ/ 54%. For Category-Goodness assimilation, /ɑ-ɔ/, /ʊ-u/ were challenging. For /ɑ-ɔ/ contrast, /ɔ/was perceived as /ɑ/ 54% of the time. For /ʊ-u/, /u/ was perceived as /ʊ/ 50% of the time. For Two-Category assimilation, the researcher decided that /ɛ-æ/ was rather challenging because /ɛ/ was perceived as /ɛ/ 57 % of the time an also perceived as /æ/ 28%. Both members of /ɪ-e^{I}/ contrast were discriminated easily.

The researcher concluded that strong assimilation of an L2 sound to an L1 category will not necessarily lead to accurate identification and discrimination. The two most accurately identified contrasts in this study were /e^I-ɛ/ and /e^I-ɪ/. For /e^I-ɛ/, participants correctly identified each sound (task 3) managing to discriminate sounds accurately (task 2) suggesting perception of these two sounds as being members of separate L2 categories. In fact, cases in which/e^I/ was identified as another vowel were rare. The researcher explained that only relying on the result of vowel assimilation task is not enough to strongly predict the level of difficulty participants experience when hearing L2 sounds. He recommended including other tasks such as identification task and discrimination task as a solution to support the findings.

As his conclusion, Jeske referred to lack of guideline for determining level of difficulty e.g. what percentage should be deemed as *easy* or *relatively easy* or *hard discrimination*. He assumed: 75% or higher can be associated with *easy discrimination*; *relatively easy* can be associated with percentages higher than the chance level all the way up to 74%; hard discrimination can be associated with percentages below the chance level.

3.5 Perceptual Similarity of North German and American English

In a study by Winfred Strange et al. (2004), the perceptual similarity of North German and American English was investigated. The achieve this, a corpus of North American English and a corpus of North German were used. As in similar previous studies, the researcher conducted perceptual assimilation tasks asking participants to categorize NG vowels they heard as their L1 (AE) categories. Participants were selected with no previous experience with the German language for the test.

To be more detailed, the aim of this study was to: (a) perceptually and acoustically compare front rounded vowels of NG to front unrounded and back rounded vowels of AE. Although in reality, 14 NG monophthongs were acoustically and perceptually compared to 11 vowels of AE, (b) investigate NG and AE mid and mid-low vowels, (c) investigate the degree to which vocalic duration may help AE listeners in the perception of L2 vowels (NG).

The hypothesis was that if long and short vowels are presented in sentence context, they would be assimilated to AE vowels more consistently compared to when presented in isolated syllable context.

In reviewing this study, I will only cover those parts relating to vowels being presented in isolated syllables. Since there were multiple tasks for acoustic and perceptual similarity of NG and AE vowels produced in citation-form syllables and in sentence context, I will only narrow down to the first assimilation task for acoustic and perceptual assimilation of NG vowels to AE vowels.

In regards to the method of acoustic similarity between NG and AE vowels, four male North German speakers who were all from northern territories in Germany were selected. They had not lived in an English speaking country during their life for more than a few months. At the time of recording the stimulus, they were all university students with their age ranging from 25 to 29. All four had studied French language for 2 to 4 years when they were at high school. The first and second author of this study were both present at the time of recoding. The models of microphone and cassette tape recorder and tape deck were mentioned**.** The participants were asked to produce 14 syllables written in German orthography on printed cards and their major errors were corrected while minor dialectal variations were ignored. Each participant produced 15 stimuli with the last utterance not included in the final corpus. After recordings were made, a North German speaker who was phonetically trained listened to all recordings and gave consent to exemplars as being good tokens of North German vowels. At the end, the corpus had 28 stimuli (2 tokens × 14 vowels) for each participant totalling 112 stimuli. The stimuli were then transferred digitally form a DAT recorder to a Macintosh computer with a sound card. For American English corpus, four native male speakers produced 11 American English vowels and the type of microphone and recorder were mentioned. The 11 American vowels were produced in hVba disyllables and each one of them was preceded by an identifying number. Finally, there were 33 stimuli (3 instances × 11 vowel). The stimuli recorded at 48 kHz were digitally transferred to a Macintosh computer through a soundcard.

In their analysis, they used SoundScope speech analysis software with the version of the software being mentioned. 112 North German stimuli and 132 of American English stimuli were displayed in Bark-scale figures for the purpose of comparison. Long and short vowels were shown by closed and open symbols respectively. The frequency for the first three formants were presented on a table in which similar vowels were shown in the same row to provide for an easier comparison.

There were spectral similarities between North German front rounded vowels and unrounded vowels. It turned out that mid long vowels had low F1 values and were high in acoustic vowel space whereas mid-high vowels such as /ɪ/, /ʏ/, /ʊ/ were higher than mid-long vowels. There were category overlaps in formant values where long and short low vowels overlapped spectrally. For American English vowels, overlapping was observed for mid-long and high-mid short vowels in F1 values. In relation to the duration, their result showed shorter duration for North German short vowels on average compared to American English vowels (NG=80 ms; AE=90 ms). On the other hand, long North German vowels were longer in duration than American counterparts (NG=153ms; AE=115 ms). To include cross-language comparisons of spectral patterns, two figures were shown in which each North German long and short vowel in F1/F2 bark space were superimposed on 11 American English vowel categories (NG vowels were circumscribed by ellipses and AE vowels were represented by dashed lines).

Then the researchers conducted a series of discriminant analysis to quantify cross-language similarity and within-language similarity in acoustic structure with F3 (third formant) included. Discriminant analysis was done in both NG and AE corpus independently with F1, F2 and F3 as input parameters for the purpose of quantification of spectral differentiation. Centres of gravity in formant values for 11 AE and NG vowels were frequency values of F1, F2 and F3. Following on from this, a second set of discriminant analysis were done by using vocalic duration as input parameter for both AE and NG corpus independently. When spectral parameter was taken into account, correct classification range for 8 NG vowels was 38% to 75% and 100% for the remaining 6 NG vowels with overall correct classification being 79%. When duration was taken into account the overall classification went remarkably higher equalling 93%. For AE corpus, overall correct classification, when only spectral parameters were involved, was 86% increasing to 96% when duration was involved as a parameter ranging from 83% to 100% across 11 vowels. In AE corpus, only /æ:/ and /ɔ:/ were not classified correctly. Their conclusion for this part was that aside from ambiguities resulting from difference between speakers, duration played an important role in differentiating NG vowels that were spectrally similar. This was not the case with AE vowels. After this, a cross language discriminant analysis was done to specify cross-language acoustic similarity in which duration was not used to solely establish spectral similarity of AE and NG. The results were presented on a table displaying AE vowels to which NG vowels were assigned to with the number of tokens also being included. The remaining tokens were presented on two different columns on a table. Their results suggested that in terms of acoustic similarity, NG front rounded vowels were somewhere between AE back rounded and front unrounded vowels. While back mid-high [ʊ] in the two AE and NG corpora was remarkably different, front mid-high [ɪ] was similar spectrally. The relative spacing of vowels in F1, F2 and F3 were different due to cross-language differences which was the reason for overlap in AE and NG corpora for low back [ɪ:] and high front [ɑ:].

Their prediction regarding perceptual similarity of NG and AE vowels stated: a) NG front rounded /y:/, /ʏ/, /ø:/, /œ/ and mid-low /e:/, /o:/, /ɛ/, /ɔ/ will be assimilated inconsistently to AE counterparts. b) NG /i:/, /ɪ/, /a:/, /u:/ will be assimilated consistently to AE counterparts.

For the method of this task, their stimulus materials were: a 112-stimulus corpus formed up the materials for the assimilation task. The speakers' production including 28 stimuli (14 vowels × 2 tokens) formed up the listening part of the task. Their listeners (participants) consisted of 12 native AE speakers who were participating in Psychology and Communication Science courses, none of whom had experience with German language nor did they have any training in a foreign language or contact with German speakers.

For this task, DAT recorder and headphone were used during the task. The 11 AE vowels were presented to listeners by words such as *hawed, hood, hid,* etc. on the screen. A response button for choosing the word that contained each vowel appeared; then, the same stimulus was presented again for the listeners to choose a number on a seven-point scale for determining the goodness of ratings. '1' meant very foreign sounding and '7' meant very English sounding. The task was conducted over the course of a few days for the purpose of familiarising the participants and better preparing them for the task. The researchers tried to divide these sessions into several parts, making it easier for the participants to meet the criterion. These sessions included several 90-minute tests in which the participants were asked to focus on different NG speakers using the entire seven-point scale when making their judgment on the stimuli.

In their result section, researchers explained the table presented including description of each columns on the table in which the frequency for most frequent answer (modal response) and second most frequent answer were mentioned. For goodness of fit, median ratings were computed with the explanation that a modal response was degree of consistency of a perceptually assimilated vowel to a certain native vowel category and median rating was the level of goodness of an NG vowel being perceived as an AE vowel category. Therefore, the percentage nearer 100% means an NG vowel is a good example of an AE category. Choosing any score near 7 on the Likert scale means that vowel is very English sounding. Accordingly, low goodness rating and lower percentage means it is very different and will be categorized as *Uncategorizable*. On a table, the perceptual assimilations were divided based on spectral similarity into 3 parts: front rounded vowels, mid and mid-low vowels, high, mid-high and low vowels.

Front rounded vowels were not assimilated to any AE vowel on a constant basis. Mid-back vowel /o:/ was assimilated to AE vowel /oʊ/ 89%, while mid-front /e:/ was assimilated to AE /i:/. NG /ʊ/ was considered as Uncategorizable, since it was not categorized as any AE vowel more than 53% of time. NG /i:/, /ɪ/, /ɑ:/, /ɛ/ were assimilated to their AE counterparts, according to researchers' prediction, with a consistency of at least 95%. Additionally, 4 other vowels /e:/, /y:/, /ʏ/, /œ/ were assimilated, to a lesser degree, to other AE counterparts with consistency between 56% to 73%. NG /i:/, /ɪ/, /ɑ:/ and /u:/ were constantly assimilated to their American counterparts.

To sum it up, their findings showed that perceptual similarity cannot always be predicted by acoustic similarity. For example, NG /o:/ and /ɛ/ were assimilated to AE categories consistently despite having acoustic differences. They explained, this may have happened due to the fact that citation-form NG and AE vowels were different in syllable structure (monosyllabic versus disyllabic). They further performed a replication of previous tasks in sentence context which I will not include in this thesis.

3.6 Overview of conclusions and the gap in previous studies

In general, their results were in line with the findings of Best (PAM-L2) and Escudero (L2LP) in terms of the extent these models can predict the level of difficulty that learners may have when perceiving L2 vowels. These studies were carried out either in the framework of PAM-L2 or L2LP only. Therefore, in this study I have endeavoured to focus on those learners whose L1 is Farsi and L2 is English. Participants are naïve learners with difficulty in understanding L2 language vowels. This study then aims to explore which vowels pose more difficulty and to what extent for the learners. It also seeks to compensate for lack of previous research (with regards to gap in previous studies) available on challenges of L2 learning by Iranian learners of English.

In case of Gilichinskaya & Strange (2010), their study showed that inexperienced Russian learners of English had exhibited similar difficulties to those immigrants with limited English. In case of Chladkova and Podlipsky (2011), their study showed that even the slightest difference in learners' native language and dialects (between-dialect difference was the term suggested by researchers) affects perception of L2 which is in line with L2LP proposed by Escudero (2005). For future cross-language studies, their suggestion was to avoid grouping participants based on their dialect differences. In case of Lengeris (2009), the study showed that having experience with duration in L1 would affect perception of L2. This study explored duration and spectral cues with researchers having used both perceptual and discrimination tasks.

In case of Jeske (2012), he referred to /æ-ɛ/ and the difficulty of learners in distinguishing /ɛ/ from /æ/. The second test in this study which was an identification test, had only 19% correct discrimination of the two vowels. This was due to Spanish speakers' lack of ability for recognising spectral difference. This strengthened the fact that studies of this type could have mixed results when it comes to PAM-L2's predictions. Jeske mentioned that despite these mixed results, it is always beneficial to add a second task (in this study it was a discrimination test) to assimilation task. In addition, he pointed out to lack of guideline in quantitatively making the level of difficulty clear as mentioned before in the designated section of this study. In case of Winfred Strange et al. (2004), their results showed that acoustic similarity cannot always predict perceptual similarity. In addition to that, participants' assimilation patterns were actually based on spectral similarity where spectral and temporal similarities conflicted.

As discussed in each of these studies, there were mixed results when it came to predicting which vowels were more difficult to perceive. Despite the possible shortcomings of assimilation tasks, PAM-L2 and L2LP are proven authentic frameworks within which cross-language studies can be carried out. Therefore, it was felt necessary to lead my own research by conducting it within the framework of PAM-L2 and L2LP.

4 Material and Methods

In terms of method of this research, the same methods which had been used in previous studies within the framework of PAM-L2 and L2LP, have been followed to estimate the extent of difficulty Iranian learners have in understanding English vowels. Hopefully, this study will help anybody interested in Farsi/Persian language to find a better understanding of how naïve learners of English perceive English vowels. For this purpose, it is necessary to know how mental representations of learners in relation to their perception of L2 sounds are suggested by PAM-L2.

To elaborate more on the concept of PAM-L2 for those readers who may not be familiar with this model providing some explanation is necessary. PAM-L2's prediction for learners' ability in discriminating L2 contrasts is based on how any member of an L2 contrast is assimilated.

Colantoni, Steele and Escudero (41) explain these assimilation patterns as follows:

Table 3. Categories of PAM-L2

Single-Category Assimilation (SC)	This refers to difficulty for learners discerning a non-native contrast as each member of an L2 vowel pair assimilates to a single L1 category. A SC contrast will be the most difficult to perceive for the participants.
Two-Category Assimilation (TC)	This indicates a scenario in which it is *easy to perceive* where each L2 sound in a contrast is assimilated to a different L1 Category. A TC contrast will be the easiest to perceive for the participants.
Category Goodness Assimilation (CG)	Both members of an L2 contrast are assimilated to an L1 Category but one is a better representative than the other which will lead to category goodness.
Uncategorized-Uncategorized (UU)	The sounds in an L2 pair may be assimilated to two or more L1 sounds.
Categorized-Uncategorized (CU)	One member of an L2 Contrast assimilates to an L1 sound, whereas the other member assimilates to two or more L1 categories. A CU contrast will be easier than single-category to perceive but more difficult compared to two-category assimilation.
Non-Assimilable (NA)	Both members of an L2 contrast have no resemblance to any L1 category.

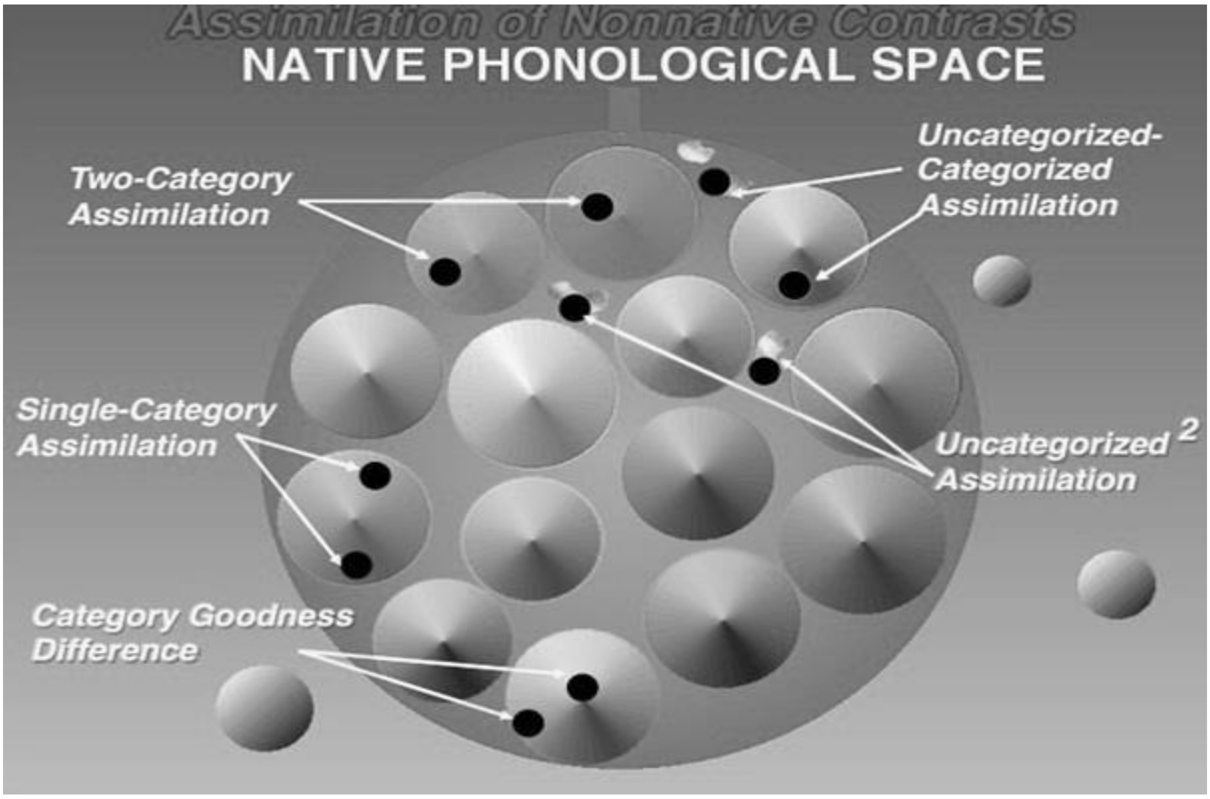


Figure 3. James E. Flege (2005) represents monolingual's assimilation of L2 contrasts to L1 phonetic categories

To make predictions within the framework of PAM-L2, 2 tests/tasks were carried out which are explained as follows:

1. Vowel Assimilation Task: It estimates how learners classify English vowels in terms of their native language vowel category.

2. Categorical Discrimination Task (AX Discrimination Task): To put this in a simple way, AX Task is Same/Different task, an old-fashioned task in an array of discrimination tasks. It aims to evaluate the results of assimilation task to see how difficult it is for learners to perceive English vowel contrasts and which vowel is indeed more difficult. There are similar tasks such as ABX, AXB, 2I2AFC, 4IAX and 4I2AFC with each one having their own set of pros and cons. The only drawback with AX task is that participants would choose the response 'different' if they are very sure of their answer.

4.1 Vowel Assimilation Task

The first test was designed by a computer-software called Praat. Praat is a Software for speech analysis and can be downloaded from the internet. ExperimentMFC (multiple forced choice) function which is a built-in feature was used. After conducting the test, results were extracted from the software.

Described below is how the first and second tests were conducted in similar fashion.

4.1.1 Methods

Participants

10 males and 10 females were selected and tested. The participants had little knowledge of English and none of them had resided in an English-speaking country except their own country Iran. Therefore, they had no prior contact or any type of interaction with Native speakers of English. Some of them had formal education up to high school level and some had only completed primary school and some had Bachelor's degree in various fields of studies; but they were nowhere close to being at an advanced level of English language learning. The purpose was to have naïve learners of English. Participants were asked about their hearing abilities and they reported no hearing difficulty or language impairments, as confirmed in the information consent forms. Some of the female participants for our vowel assimilation task were housewives, others were social workers or unemployed. As for the male participants, some of them were pensioners while others were working full-time jobs. There were also unemployed participants in the test. The *mean* for age of participants was 44 with *standard deviation* being 14.

Stimuli

A part of stimuli from the study by Pereira (2014) were used. In that study she had four native speakers of Southern British English (2 males & 2 females) who recorded a list of 61 randomized words containing English vowels in /bVt/ and /hVd/ contexts. Three repetitions per word were included for a selection to be made later. A Bruel & Kjaer 2231 microphone was connected to a DAT recorder and recordings were made at a sampling rate of 48 kHz. They had later chosen best sample for each vowel in each context. In this study, we only used recordings made in /hVd/ contexts (11 vowels x 4 speakers).

Procedure

The participants were instructed to listen to the SSBE words which contained target vowels and then were asked to choose which FA vowel sounded most similar to that token. The words selected for these two tests were both nonsense words such as 'hudd' and meaningful words such as 'hard'. The target vowels were embedded in these meaningful/nonsense words in /hVd/

contexts. After a FA vowel category was selected, they were asked to make their decision on the perceived goodness of them on a nine-point Likert scale (1=unlike to 9=identical). Three repetitions of the 44 SSBE stimuli were played in a randomized order over high-quality Beyerdynamic DT 231 Headphones had been borrowed from UEF and we used my personal ASUS Laptop set at a comfortable listening level. Each trial consisted of a vowel token that listeners identified by selecting on the computer screen from one of the six FA vowel response options. These options written in FA orthography, in separate response boxes as shown below at the bottom of this section, to give readers an idea of what our participants faced upon taking the test. Participants were given the chance to hear the SSBE several times by clicking the designed replay button, before selecting the appropriate response.

At the beginning of the test, detailed verbal explanation had been given to our participants in Farsi (FA) language, on what was expected of them when the test would begin. Since they had little knowledge of English, I had to give examples in Farsi language. There were repetitions of 11 vowels in /hVd/ context, for a total of 88 trials. The test was run in Praat software.

اَ اِ اُ آ او ای

Above are Farsi vowels written in Farsi orthography. They were shown to participants in exactly the same style using Praat. Participants were familiar with Farsi orthography because all of them had primary school education at least and were all familiar with English orthography due to their secondary school education background. Therefore, they had no problem with assimilation and identification tasks in terms of reading from screen. Participants were all literate enough for the purpose of this study.

4.2 Categorical discrimination task

For identification task, they had no problem in classifying target vowels as same or different. They were given instruction on how to choose and were familiar with English and Farsi orthography which enabled them to select their responses.

4.2.1 Methods

Participants

All 20 participants of our assimilation task performed this task right after completing assimilation task to see how they could establish mental representation of English vowels.

Stimuli

The stimuli in the /hVd/ context were presented to the participants through the same Headphones and the same laptop that were used for the first test.

Procedure

As explained earlier, this test was done by participants right after they completed the first task. We also had to consider that due to high number of repetitions participants could get tired or bored increasing carelessness or risk of making error. The participants heard two words, then asked to choose if they were different or the same. There were repetitions of the same 11 vowels for a total of 116 trials. The test was designed and run in Praat software using ExperimentMFC function.

5 Linear discriminant analysis

5.1 Methods

The following sections will provide information on methods through which SSBE and Farsi vowel formants have been extracted and used in the present study. A Linear Discriminant Analysis has been used to investigate to what extent there would be correct classifications in accurate distinction of six Farsi vowels and eleven SSBE vowels. The intent here was to look, at the issue of how distinct Farsi vowels can be kept in the auditory perception. Wang & Van Heuven (2006) mentioned "if the categories in dataset are more distinct then the number of classification errors yielded by the algorithm will be fewer".

5.1.1 Vowel formants data used for acoustic similarity analyses

The SSBE vowel formant data is compiled from a study by Hawkins and Midgley (2005). The SSBE data used here is based on productions of ten male native British English subjects whose ages ranged from 25-40 years. They were all students or staff of the University of Cambridge. The Farsi vowel formant data is compiled from a study by Ghaffarvand Mokari & Werner (2017) for male participants.

According to Ghaffarvand Mokari and Werner (2017), the participants selected for recording Farsi vowels had volunteered for the task. Therefore, they were asked to put their signature on informed consent forms distributed before they entered the study. There were background questionnaires that had to be completed in order for the actual recordings to take place. There was also a criteria that had to be met by volunteers/participants; in other words they were deemed as being eligible for taking part in the study only if they met the following criteria: (a) being a monolingual speaker of Farsi (b) being born and raised in Tehran which is the capital of Iran and having used the Tehrani dialect in daily communications on a day to day basis (c) being aged 20–40 years old (d) being a university student/graduated student with no prior speech/language disorder and never dealing with similar disorders. There was a fixed carrier sentence in which six Farsi vowels were embedded with the same stress pattern (‘vɒdʒeje Y mɒnænde X æst’ meaning ‘the word Y is similar to the word X’) where the ‘X’ was the target word and ‘Y’ was a rhyming word. To give the participants a clue about pronunciation of the coming target vowels, they placed a rhyming real CVC word (Y). Most of the target words were real and meaningful terms. It is necessary to mention that the ‘X’ was always in an accented position in a sentence where sentences were introduced to the speakers on printed lists. Every target vowel was produced for a total number of three times in three different target words in a carrier sentence. It was demanded of the participants to read the list of carrier sentences for a total number of three times, because it was decided it was necessary to drop misproduced and creaky tokens excluding them from the final recordings. They then had at least 2815 tokens waiting to be analysed. When the actual recording began, utterances were recorded in digital form in a sound-attenuated booth, using a Sennheiser EW112-P G3 microphone connected to a Zoom H4N recorder.

In order to provide more detailed information regarding the distance between speaker’s mouth and microphone, it can be confirmed this was set at 18–20 cm. The sampling rate of recordings was 44100 Hz with a 16-bit resolution. The recordings were segmented and labelled manually by the use of Praat speech analysis software with the given version being 6.0.06 (Boersma & Weenink, 2015). Formant frequency analysis was performed using an algorithm by the name of Burg which is available in Praat by default. Burg algorithm has standard settings of its own for females (range of 5500 Hz for five formants) and also for males (range of 5000 Hz for five formants). In terms of time interval, the set up was fixed at 0.005 seconds for formant frequency samplings. Formant trackings were checked to find errors that could potentially be made. In case an error was found, it was obliged to correct it and so it was manually modified.

Ghaffarvand Mokari and Werner mentioned “the first three formants (F1, F2, F3) were calculated by averaging the values at central 40% of vowel duration for each token. The reason for this was to avoid sole reliance on one point with higher risk of unreal measures”.

The data for SSBE and Farsi were gathered in a similar manner. All recordings were conducted in quiet rooms using high quality microphones. The target vowels were recorded in /cVc/ context. For SSBE vowels, each of the eleven monophthongal vowels was recorded in /hVd/ context. Farsi vowels that were used were recorded tokens in a /hVd/ consonantal context including male participants. According to Hawkins and Midgleley (2005), “for SSBE vowels the target words which included SSBE vowels along with nineteen monosyllabic filler words were randomized in two independent lists. Two further lists were made by reversing the order of words in the first two lists for a total of four repetitions per word. Each list began and ended with at least one filler word for the purpose of avoiding *beginning- and end-of-list effects* in reading” (186). Also for having natural tokens of Farsi vowels, the target words were recorded in a fixed carrier phrase. In both studies, for every single vowel, close to middle of the vowels and at their steady state, frequencies for F1 and F2 were measured. Four repetitions of SSBE vowels and three repetitions of Farsi vowels were recorded. For the SSBE vowels, the digitization of the test words took place on a Silicon Graphics computer. For this purpose, a 16 kHz sampling rate was used and again frequencies for F1 and F2, close to middle of vowels and at their steady state were measured whenever possible. According to Hawkins and Midgeley (2005), “measurements were made with the xwaves + formant tracker, using 30 ms, 18-pole autocorrelation spectra with a cos**4 window. The step-size for successive spectra was 30 ms. Missed formants were measured manually using both Discrete Fourier Transform (Hanning window) and Burg linear predictive coding (rectangular window) spectra, with a window of 50 ms” (187).

5.1.2 Vowel contrasts

After comparing the acoustic and perceptual assimilation patterns of L1 to L2 vowels individually, we will consider the predictions for discrimination of different vowel contrasts. We will compare the predictions based on acoustic and perceptual assimilations for eight SSBE vowel contrasts including /ʊ-u/, /ɑ-ʌ/, /ɔ-ɒ/, /æ-ʌ/, /æ-ɜ/, /i-ɪ/, /i-e/ and /ɔ-u/. These vowel contrasts were chosen based on their *adjacency in the SSBE vowel system* and their *possible difficulty* for Farsi listeners.

6 Results

6.1 Perceptual assimilation of SSBE to FA vowels

This section first provides the results from perceptual assimilation task, PAM and L2LP's predictions for eight SSBE contrasts' discriminations based on our results. Equally, the results of discrimination task are also provided which will be compared to the predictions made by vowel assimilation task.

Table 4 displays percentage of times a FA vowel was selected for each of the 11 SSBE vowels using summed responses across all 20 listeners. This table only shows those FA vowel responses that were used by listeners at least 3% of the time.

Table 4. Perceptual assimilation results for categorization of SSBE vowels as FA vowels. The goodness ratings (GR) and standard deviations (SD) are shown next to the mean percentage of each vowel. The goodness ratings (GR) were in a scale of 1-9.

SSBE vowels	**FA vowels**											
	i	**GR (SD)**	**e**	**GR (SD)**	**æ**	**GR (SD)**	**u**	**GR (SD)**	**o**	**GR (SD)**	**ɒ**	**GR (SD)**
i	**80.62**	7.3 (1.9)	16.87	6.7 (1.4)								
ɪ	**38.75**	6.5 (2.1)	**57.5**	6.9 (1.8)								
ɛ	3.12	5 (1.5)	**70**	6.7 (2)	**25**	7.3 (1.4)						
æ			8.12	6.6 (1.7)	**85**	6.9 (2)					5.625	5.5 (1.6)
ɜ	6.25	5.8 (1.7)	16.25	6.5 (1.7)	**40.62**	6.6 (2)	6.87	6 (2.2)	11.87	6.4 (1.7)	18.12	6.6 (1.7)
ʌ					**44.37**	6.3 (1.9)			3.12	5.6 (1.9)	**46.25**	6.7 (2)
u	19.37	7 (1.4)	3.75	5.5 (1.3)			**50**	6.8 (2.1)	**23.75**	6.7 (1.8)		
ʊ	5	6.3 (1.6)	**21.25**	6.4 (2)			**43.12**	6.7 (1.9)	**28.12**	6.7 (1.8)		
ɔ							**30.62**	6.6 (2)	**66.87**	7.1 (1.8)		
ɒ					3.12	6.8 (2.1)	14.37	6.1 (1.4)	**56.25**	6.7 (1.9)	**24.37**	7.2 (1.8)
ɑ					11.87	6.1 (1.4)					**80.62**	7.2 (1.8)

As it can be seen in Table 4, the SSBE /i/, /æ/ and /ɑ/ vowels were overwhelmingly (>80%) assimilated to FA vowels /i/, /æ/ and /ɒ/, respectively. The SSBE /i/ was rated with a GR of 7.3 to its FA counterpart. In a similar fashion, assimilation of SSBE /ɑ/ and /æ/ received GRs of 7.2

and 6.9 to their FA counterparts, respectively. The SSBE /ɛ/ was categorized as FA /e/ 70% (GR = 6.7) of the time. The SSBE /ɔ/ and /ɒ/ vowels were categorized as FA /o/ 66.87% and 56.25% of the time, respectively. The SSBE /ɪ/ was assimilated to the FA /e/ with 57.5% categorization rate. The SSBE high-back rounded /u/ was assimilated to FA counterpart vowel /u/ 50% of the time.

Other SSBE vowels were not assimilated to any FA vowel more than 50% of the time. The SSBE /ʌ/ was mostly categorized either as FA /æ/ or /ɒ/. Interestingly, The SSBE /ɜ/ was assimilated to all FA vowels with different percentage rates, but the highest assimilation rate for SSBE /ɜ/ was to FA /æ/ (40.62%).

Table 5. Assimilation percentages and goodness ratings (in a scale of 1-9) of SSBE vowels to FA vowels and the assimilation patterns/scenarios based on PAM and L2LP with a 50% cut-off criterion.

L2 vowel contrast	Assimilation to L1	Assimilation Pattern (PAM)	Assimilation Scenario (L2LP)
/ʊ-u/	/ʊ/ → /u/ (43.12%, GR: 6.7) /u/ → /u/ (50%, GR: 6.8)	uncategorized-uncategorized assimilation	SUBSET
/ɑ-ʌ/	/ɑ/ → /ɒ/ (80.62%, GR: 7.2) /ʌ/ → /ɒ/ (46.25%, GR: 6.7)	categorized-uncategorized	SUBSET
/ɔ-ɒ/	/ɔ/ → /o/ (66.87%, GR: 7.1) /ɒ/ → /o/ (56.25%, GR: 6.7)	single-category assimilation	NEW
/æ-ʌ/	/æ/ → /æ/ (85%, GR: 6.9) /ʌ/ → /ɒ/ (46.25%, GR: 6.7)	categorized-uncategorized	SUBSET
/æ-ɜ/	/æ/ → /æ/ (85%, GR: 6.9) /ɜ/ → /y/ (40.62%, GR: 6.6)	categorized-uncategorized	SUBSET
/i-ɪ/	/i/ → /i/ (80.62%, GR: 7.3) /ɪ/ → /e/ (57.5%, GR: 6.9)	two-category assimilation	SIMILAR
/i-ɛ/	/i/ → /i/ (80.62%, GR: 7.3) /ɛ/ → /e/ (70%, GR: 6.7)	two-category assimilation	SIMILAR
/ɔ-u/	/ɔ/ → /o/ (66.87%, GR: 7.1) /u/ → /u/ (50%, GR: 6.8)	two-category assimilation	SIMILAR
/æ-ɑ/	/æ/ → /æ/ (85%, GR: 6.9) /ɑ/ → /ɒ/ (80.62%, GR: 7.2)	two-category assimilation	SIMILAR

Based on perceptual assimilations, Table 4 displays assimilation patterns/scenarios within the framework of PAM and L2LP models. As mentioned earlier in the methods section, these models' predictions would be put into test for nine SSBE vowel contrasts including /ʊ-u/, /ɑ-ʌ/, /ɔ-ɒ/, /æ-ʌ/, /æ-ɜ/, /i-ɪ/, /i-ɛ/, /ɔ-u/ and /æ-ɑ/.

6.2 Categorical discrimination task

The Sensitivity A' Score (Sondgrass et al. 1985) was calculated for each contrast discrimination from pre/post-test results. It is necessary to mention that this calculation was actually based on proportion of hits and false alarms for each contrast. To reduce likely effects of response bias as much as possible, A' score was calculated. In case of selecting the odd item out in change trials correctly, the definition for it is hits. Similarly, in case of selecting an odd item out in no-change trials incorrectly, the definition for it is false alarms. A' score of 1.000 is an indication of existence of perfect sensitivity to a vowel pair (i.e. correct responses to all change and no-change trials). A' score of 0.500 is an indication of existence of a theoretically defined chance level, which by itself means that discrimination is by chance and there is lack of phonetic sensitivity (Rallo Fabra and Romero, 2012).

Figure 4 and Table 5 show the A' mean scores (SD) for discrimination of the nine vowel contrasts. As seen in Figure 3 and Table 3, the most difficult vowel contrast to be discriminated was /ʊ-u/ with A' score of 0.675 followed by /ɑ-ʌ/, /ɔ-ɒ/, /æ-ʌ/, /æ-ɜ/, /i-ɪ/, /i-e/ and /ɔ-u/.

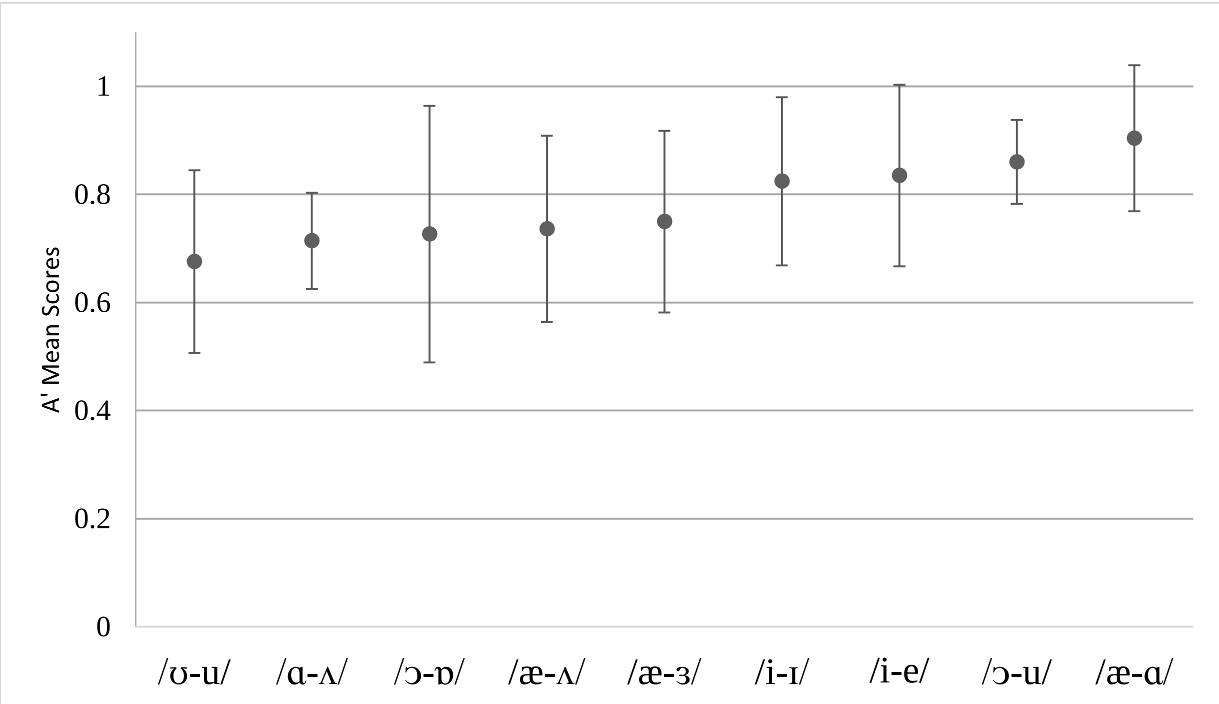


Figure 4. The mean A' scores for 9 SSBE contrasts discriminated by FA learners.

As explained above, A' scores of 0.6 to 1.00 suggest phonetic sensitivity of the participants and values of 0.5 or less refer to a discrimination by chance with no phonetic sensitivity (Rallo Fabra et al., 2012). Therefore, it can be assumed that there was no vowel contrast which FA listeners were not at all sensitive to. On the other hand, there were differences between the sensitivity of them ranging from A' scores of 0.675 to 0.903.

Table 5. presents A' prime scores and Standard deviations calculated for each contrast.

Vowel contrast	A' prime	Standard deviation
/ʊ-u/	0.675	0.169
/ɑ-ʌ/	0.714	0.089
/ɔ-ɒ/	0.726	0.237
/æ-ʌ/	0.736	0.172
/æ-ɜ/	0.749	0.167
/i-ɪ/	0.824	0.155
/i- ɛ/	0.834	0.167
/ɔ-u/	0.859	0.077
/æ-ɑ/	0.903	0.135

A one-way analysis of variance (ANOVA) showed a notable difference between discrimination scores of the vowel pairs [$F (8, 160) = 4.973$, $p < 0.001$].

6.3 Linear discriminant analysis

In this section, first the LDA (Linear Discriminant Anlaysis) results are reported which was used to establish the quantitative spectral similarity of SSBE and FA vowels. Then, the predicted categorization patterns of each SSBE vowel pair based on the LDA and according to L2LP model is provided. At last, the previously found perceptual assimilation results are represented and compared to LDA results.

For the purpose of quantitatively comparing the vowel systems FA and SSBE, we conducted a series of LDAs. According to Nishi et al. (2008), “discriminant analysis is a multidimensional correlational technique that establishes classification rules to maximally distinguish two or more predefined nominal categories from one another i.e. vowel categories. The classification rules are specified by linear combinations of input variables i.e. acoustic measures such as formant frequencies and vowel duration with weights for each variable determined to maximize separation of categories using the values for centre of gravity of each category and within-category dispersion in the input set. By applying the classification rule established for the input set, membership of tokens in the input set can be re-evaluated posterior classification technique and the amount of overlap between categories can be quantified in terms of incorrect classification of tokens. It is also possible to classify a new set of data test set using the previously established classification rules” (p. 578).

A separate LDA was conducted for each vowel using the F1/ F2 (in Bark) of the FA and SSBE vowel tokens as described earlier in the methods section. In Table 6 and Table 7, the discriminant functions are based on linear combinations of the two spectral parameters, F1 and F2, for FA and SSBE.

Table 6. Classification of Farsi vowels produced in hVd context (percentage) based on F1 and F2 Bark values for male speakers.

	ɒ	**æ**	**e**	**i**	**o**	**u**
ɒ	97.62	0	0	0	1.19	0
æ	0	98.81	0	0	0	0
e	0	1.19	100	0	0	0
i	0	0	0	100	0	0
o	2.38	0	0	0	98.81	2.38
u	0	0	0	0	0	97.62

Table 7. Classification of SSBE vowels (percentage) based on F1 and F2 Bark values for male speakers.

	æ	**ɑ**	**ɛ**	**i**	**ɜ**	**ɪ**	**ɔ**	**ɒ**	**ʊ**	**ʌ**	**u**
æ	80	0	2.5	0	0	0	0	0	0	5	0
ɑ	0	72.5	0	0	0	0	0	2.5	0	15	0
ɛ	5	0	82.5	0	2.5	2.5	0	0	0	0	0
i	0	0	0	92.5	0	0	0	0	0	0	2.5
ɜ	15	0	0	0	95	0	0	0	20	15	0
ɪ	0	0	15	7.5	0	97.5	0	0	2.5	0	15
ɔ	0	0	0	0	0	0	95	2.5	7.5	0	0
ɒ	0	7.5	0	0	0	0	5	95	15	0	0
ʊ	0	5	0	0	2.5	0	0	0	47.5	5	12.5
ʌ	0	15	0	0	0	0	0	0	0	60	0
u	0	0	0	0	0	0	0	0	7.5	0	70

Based on formant values, the model for FA yielded 98.81% correct classification for the training vowels and the model for SSBE yielded 80.68% correct classification for trained tokens.

A cross-language LDA was conducted for establishing the spectral similarity of SSBE and FA. There is corresponding similarity between SSBE and FA vowels which can be shown in a context. Therefore, the stimuli of the SSBE corpus serving as the test corpus were classified with regards to FA vowel centres of gravity established for SSBE input corpus (Strange et al., 2004). Accordingly, the discriminant analysis will show to what extent SSBE tokens will fit with centres of gravity of the input corpus tokens (SSBE), providing percentages that every vowel is categorized as one of the FA vowel contrasts. As a result, any FA vowel contrast which will have highest percentage (highest compatibility with an SSBE counterpart) for an SSBE vowel, is indicator of being acoustically closest to the non-native vowel. Table 8 shows the LDA results for SSBE vowels classified by FA vowel.

Table 8. LDA classification results (percentages) for SSBE tokens based on trained model for FA vowels.

		SSBE vowels										
		æ	ɑ	ɛ	i	ɜ	ɪ	ɔ	ɒ	ʊ	ʌ	u
FA vowels	ɒ	-	85	-	-	12.5	-	-	12.5	-	60	-
	æ	95	-	20	-	17.5	-	-	-	-	25	-
	e	5	-	80	-	55	30	-	-	22.5	-	32.5
	i	-	-	-	100	-	70	-	-	-	-	42.5
	o	-	15	-	-	15	-	35	85	52.5	15	-
	u	-	-	-	-	-	-	65	2.5	25	-	25

For the SSBE /æ/, the model classified the closest FA /æ/ vowel 95% of the time. A similar case was found for some other SSBE vowels such as the SSBE /ɑ/, /ɛ/, /i/, /ɪ/ and /ɒ/ which were classified as FA /ɑ/ 85%, as FA /e/ 80%, as FA /i/ 100%, as FA /i/ 70% and as FA /o/ 85% of the time, respectively. For some of the SSBE vowels, the model classified the tokens across a range of FA vowels. The SSBE /ɜ/ was classified as FA /e/ 55% and FA /æ/ 17.5%. SSBE /ɔ/ was classified as FA /u/ 65% and FA /o/ 35%. SSBE /ʊ/ was classified as FA /o/ with 52.5%. SSBE /ʌ/ classified as FA /ɒ/ and /æ/ with 60% and 25%, respectively. Finally, SSBE /u/ was classified as FA /i/ and /e/ and /u/ with 42.5%, 32.5% and 25%, respectively.

6.4 Predictions for discrimination of L2 contrasts

Figure 5 presents the predicted categorization patterns for each SSBE contrast based on the LDA. Categorization pattern from LDA results were used for predicting perceptual discriminations.

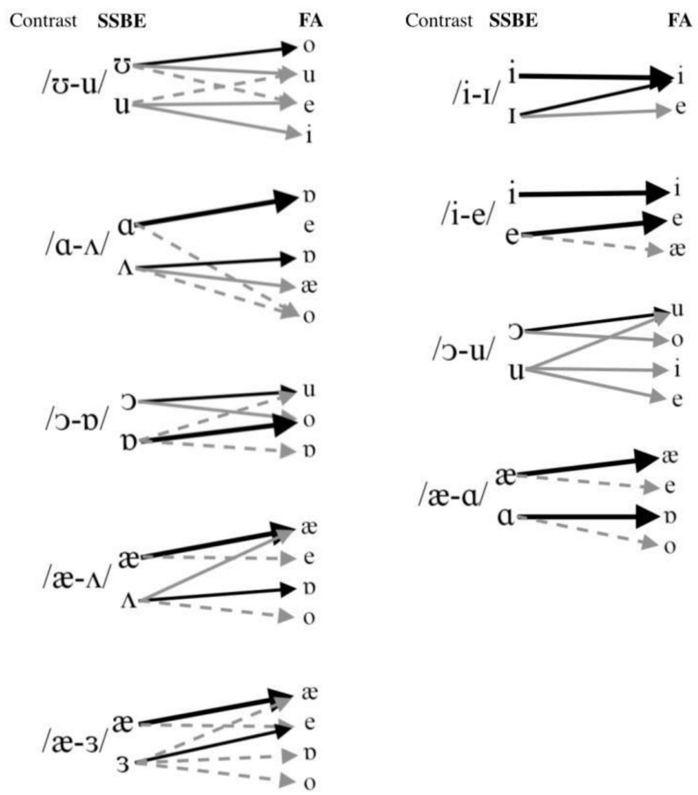


Figure 5. FA non-native categorization patterns based on the percentage of categorizations in LDA results. Percentages > 75% = thick black line, between 50% and 75% = thinner black line, between 25% and 50% = thin grey line and between 0.08% and 25% = thin dotted grey line.

To have a quantitative approach in order to enable predicting L2 vowel discrimination difficulty following Levy (2009) and Elvin (2016) and Ghaffarvand Mokari (2018), a *cross-language assimilation overlap* method has been used based on cross-linguistic LDA results and perceptual assimilation results which has rendered a score of overlap between each member of L2 contrasts and L1 vowels (Table 9 and Table 10).

Table 9. Acoustic overlap scores for SSBE vowel contrasts.

SSBE contrasts	Overlap	SSBE contrasts	Overlap
/ʊ-u/	47.5	/i-ɪ/	70.0
/ɑ-ʌ/	75.0	/i-ɛ/	0.0
/ɔ-ɒ/	15.0	/ɔ-u/	25.0
/æ-ʌ/	25.0	/æ-ɑ/	0.0
/æ-ɜ/	22.5		

Table 10. Perceptual overlap of SSBE contrasts for FA listeners.

SSBE contrasts	Overlap	SSBE contrasts	Overlap
/ʊ-u/	75.6	/i-ɪ/	55.6
/ɑ-ʌ/	58.1	/i-ɛ/	20.0
/ɔ-ɒ/	70.6	/ɔ-u/	54.4
/æ-ʌ/	50.0	/æ-ɑ/	17.5
/æ-ɜ/	54.4		

As explained before these tables (table 9 and 10) provide information on the result of cross-language assimilation overlap calculation which is about the overlap scores between members of an L2 contrast and L1 vowels. Table 9 shows acoustic overlap scores and table 10 shows perceptual overlap scores of SSBE contrasts for FA listeners.

Table 11. Predictions within the framework of L2LP based on the acoustic similarities of SSBE vowels to FA vowels.

L2 vowel contrast	**Acoustic Assimilation to L1**	**Assimilation Scenarios (L2LP)**
/ʊ-u/	/ʊ/ → /o/ (52.5%) /u/ → /i/ (42.5%)	SUBSET
/ɑ-ʌ/	/ɑ/ → /ɒ/ (85%) /ʌ/ → /ɒ/ (37.5%)	SUBSET
/ɔ-ɒ/	/ɔ/ → /u/ (65%) /ɒ/ → /o/ (85%)	SIMILAR
/æ-ʌ/	/æ/ → /æ/ (95%) /ʌ/ → /ɒ/ (60%)	SIMILAR
/æ-ɜ/	/æ/ → / æ/ (95%) /ɜ/ → /e/ (55%)	SIMILAR
/i-ɪ/	/i/ → /i/ (100%) /ɪ/ → /i/ (70%)	NEW
/i-ɛ /	/i/ → /i/ (100%) /ɛ/ → /e/ (80%)	SIMILAR
/ɔ-u/	/ɔ/ → /u/ (65%) /u/ → /i/ (42.5%)	SIMILAR
/æ-ɑ/	/æ/ → /æ/ (95%) /ɑ/ → /ɒ/ (85%)	SIMILAR

As a summation for acoustic assimilations within the framework of L2LP, predictions made within this framework are shown on this table which will be compared to perceptual assimilations represented on table 12.

Table 12. Acoustic & Perceptual overlap scores for SSBE vowel contrasts

SSBE contrasts	Acoustic Overlap	SSBE contrasts	Perceptual Overlap
/ʊ-u/	47.5	/ʊ-u/	75.6
/ɑ-ʌ/	75.0	/ɑ-ʌ/	58.1
/ɔ-ɒ/	15.0	/ɔ-ɒ/	70.6
/æ-ʌ/	25.0	/æ-ʌ/	50.0
/æ-ɜ/	22.5	/æ-ɜ/	54.4
/i-ɪ/	70.0	/i-ɪ/	55.6
/i-ɛ/	0.0	/i-ɛ/	20.0
/ɔ-u/	25.0	/ɔ-u/	54.4
/æ-ɑ/	0.0	/æ-ɑ/	17.5

On this table, acoustic and perceptual overlap scores have been presented side by side which makes it visually easier to compare.

Table 13. shows predictions based on both acoustic and perceptual similarities of SSBE vowels to FA vowels.

L2 vowel contrast	**Acoustic Assimilation to L1**	**Assimilation Scenarios (L2LP)**	**Perceptual Assimilation to L1**	**Assimilation Scenario (L2LP)**
/ʊ-u/	/ʊ/ → /o/ (52.5%) /u/ → /i/ (42.5%)	SUBSET	/ʊ/ → /u/ (43.12%) /u/ → /u/ (50%)	SUBSET
/ɑ-ʌ/	/ɑ/ → /ɒ/ (85%) /ʌ/ → /ɒ/ (37.5%)	SUBSET	/ɑ/ → /ɒ/ (80.62%) /ʌ/ → /ɒ/ (46.25%)	SUBSET
/ɔ-ɒ/	/ɔ/ → /u/ (65%) /ɒ/ → /o/ (85%)	SIMILAR	/ɔ/ → /o/ (66.87%) /ɒ/ → /o/ (56.25%)	NEW
/æ-ʌ/	/æ/ → /æ/ (95%) /ʌ/ → /ɒ/ (60%)	SIMILAR	/æ/ → /æ/ (85%) /ʌ/ → /ɒ/ (46.25%)	SUBSET
/æ-ɜ/	/æ/ → / æ/ (95%) /ɜ/ → /e/ (55%)	SIMILAR	/æ/ → /æ/ (85%) /ɜ/ → /y/ (40.62%)	SUBSET
/i-ɪ/	/i/ → /i/ (100%) /ɪ/ → /i/ (70%)	NEW	/i/ → /i/ (80.62%) /ɪ/ → /e/ (57.5%)	SIMILAR
/i-ɛ/	/i/ → /i/ (100%) /ɛ/ → /e/ (80%)	SIMILAR	/i/ → /i/ (80.62%) /ɛ/ → /e/ (70%)	SIMILAR
/ɔ-u/	/ɔ/ → /u/ (65%) /u/ → /i/ (42.5%)	SIMILAR	/ɔ/ → /o/ (66.87%) /u/ → /u/ (50%)	SIMILAR
/æ-ɑ/	/æ/ → /æ/ (95%) /ɑ/ → /ɒ/ (85%)	SIMILAR	/æ/ → /æ/ (85%) /ɑ/ → /ɒ/ (80.62%)	SIMILAR

On this table we can see acoustic assimilation to L1 and assimilation scenarios based on L2LP as well as perceptual assimilations to L1 and their respective scenarios. For four contrasts, the predictions of acoustic and perceptual assimilation are different which have been discussed in details before. As can be seen, there are cases of SUBSET, NEW and SIMILAR scenarios for both perceptual and acoustic assimilations based on L2LP model.

7 Discussion

In this section, we will firstly discuss the predictions made for eight SSBE vowel contrasts in the context of PAM and L2LP models. Secondly, the results of the discrimination task will be discussed and compared to the predictions made possible within the framework of PAM and L2LP.

In /ʊ-u/ contrast, both members of contrast were assimilated mostly to FA /u/. However, the SSBE /u/ was in borderline of cut-off value (50%) whereas /ʊ/ was assimilated to FA /u/ less than 50% of the time. If we assume the SSBE /u/ is categorized as the FA /u/, this contrast would be a case of "categorized-uncategorized" ("subset scenario" in L2LP terms). The prediction of PAM and L2LP for this contrast is that it is less difficult than SC / "new scenario" to discriminate by non-native (naïve) listeners. On the other hand, if the SSBE /u/ is regarded as "uncategorized" then this contrast is expected to be a case of UU assimilation with prediction for discrimination of this contrast varying from poor to excellent, depending on phonetic similarity of two phones to one another and also to native phonological categories. In this case, our assumption for discrimination of this contrast would not be accurate, as both members of this contrast are assimilated to FA /u/ with near cut-off values and similar goodness ratings.

The /æ-ʌ/ and /æ-ɜ/ contrasts can be regarded as being cases of UC / SUBSET because one member of each contrast is not categorized as any FA vowel more than 50% of the time. The SSBE /ʌ/ was assimilated to either FA /æ/ or /ɒ/ with 44.37% and 46.25% assimilation percentages, respectively. In a similar fashion, the SSBE /ɜ/ was assimilated to different FA vowels with different rates. Therefore, the difficulty of discrimination of these two contrasts is also predicted to be easier than SC cases.

The SSBE /ɔ-ɒ/ contrast can be regarded as a case of SC / NEW, due to the fact that both members of this contrast are mostly assimilated as single FA /o/ category. The SSBE contrasts

/i-ɪ/, /i-ɛ/, /ɔ-u/ and /æ-ɑ/ are regarded as cases of TC / SIMILAR assimilations, because each vowel in these contrasts is assimilated to different FA vowel categories. Both PAM and L2LP predict these pairs as the easiest ones to discriminate by non-native listeners.

The results of discrimination task revealed that all nine vowel contrasts were discriminated with scores above chance level. However, there were differences regarding the difficulty of these contrasts as suggested by discrimination scores and ANOVA results. The results of discrimination task were mostly in line with predictions of PAM and L2LP models.

The SSBE /æ/ and /ɑ/ were assimilated to FA /æ/ and /ɒ/, respectively. Their GR scores show that they were both identified as similar tokens compared to their FA counterparts (/æ/ → /æ/: GR = 6.9 and /ɑ/ → /ɒ/: GR = 7.2). According to PAM and L2LP, the SSBE /æ-ɑ/ is a case of TC / SIMILAR assimilation/scenario thus predicted to be easily discriminated by FA listeners. This was in line with discrimination results as this contrast was discriminated by FA listeners with more than 0.9 sensitivity score. Similarly, /i-ɪ/, /i-ɛ/ and /ɔ-u/ contrasts as the cases of TC / SIMILAR assimilations/scenarios were predicted to be discriminated very easily. The discrimination results also revealed that these contrasts were also very easy for FA listeners to discriminate (with > 0.8 sensitivity scores).

On the other hand, /ɑ-ʌ/, /æ-ʌ/ and /æ-ɜ/ were also predicted to be (moderately) easy in discrimination for FA listeners, since they were classified as cases of UC / SUBSET. Their reception of good sensitivity scores, although less than the TC / SIMILAR cases, was in line with the prediction of both PAM and L2LP models.

Among the eight contrasts put into test, the SSBE /ʊ-u/ was the most difficult contrast for FA listeners (although still with sensitivity score > 6). The SSBE /ʊ/ and /u/ were assimilated to FA /u/ with similar assimilation percentages. As we mentioned earlier, if we consider this case as a UU / SUBSET assimilation/scenario, PAM predicts that discrimination of UU contrast will be in a range from poor to excellent depending on phonetic similarity of two phones to one another and to native phonological categories. In line with PAM predictions, since both members of the pair are assimilated to a similar FA category even with similar percentages and similar GRs, it can be expected to see a moderately difficult discrimination with slightly better discrimination score than a case of poor discrimination.

Finally, /ɔ-ɒ/ was expected to fall into a SC / NEW assimilation/scenario since both members are assimilated to a single category in FA language (/o/). Based on both PAM and L2LP, we predict this pair to be a very difficult contrast to discriminate by FA listeners. Contrary to this prediction, the results of the discrimination task revealed it to be moderately easy to discriminate by FA listeners.

In this study, the predictions for discrimination of SSBE vowels by Farsi speakers were investigated based on perceptual assimilation results within the framework of PAM-L2 and L2LP. Acoustic properties were then compared to predict discrimination of SSBE vowels.

In general, the results of linear discriminant analysis were to some degree consistent with assimilation results which will be discussed further. Based on our results we can say the findings of this study were in line with findings of Ghaffarvand Mokari (2018), suggesting that Linear Discriminant analysis cannot precisely predict perceptual assimilation for all contrasts. Perceptual assimilation task shows interesting information about those vowels which are difficult for Farsi learners of English. For example, SSBE /u, ʊ, ɔ/ were assimilated as FA /u/ which is a demonstration that these tense/lax vowels in English language pose difficulty for FA learners with limited vowel inventory and limited exposure to L2 language. The results of vowel assimilations task were further put into another test with an identification task following methods of previous studies. As in previous studies, it once again proves that perceptual assimilation task cannot always be accurate even though it is good measure for cross-language investigation.

Following Elvin (2016) and Ghaffarvand Mokari (2018), we would predict that those contrast that have little or no overlap scores could be easier to discriminate than those contrasts with higher overlap scores. Therefore, based on perceptual overlap scores and categorization patterns based on LDA results, the predictions for /æ-ʌ/, /æ-ɜ/ may range from moderately difficult to relatively easy. It suggests that /i-ɛ/ and /æ-ɑ/ contrasts are relatively easy and /ɔ-u/ can be moderately difficult and /i-ɪ/ can be difficult for Farsi listeners. Other contrasts did not prove the same overlap patterns.

The results of the assimilation task showed that SSBE /ɛ/, /æ/, /ɔ/, /ɒ/ were categorized as FA /e/, /e/, /æ/, /o/, /o/ for at least 50% of the time. Additionally, the result of acoustic task showed that SSBE /ɑ/, /i/, /ɪ/, /ɔ/, /ɒ/, /ʊ/, /ʌ/ were categorized as FA /ɒ/, /i/, /i/, /u/, /o/, /o/, /ɒ/ for at least 50% of the time and more. Therefore, by comparing the perceptual assimilation results with those of LDA classifications, it is seen that SSBE /i/, /ɛ/, /æ/, /ɑ/, /ɒ/ were perceptually categorized as single native categories FA /i/, /e/, /æ/, /ɒ/, /o/ for more than 50 % of the time, respectively.

The results of this study showed that LDA cross-language classifications based on acoustic properties of vowels could not precisely predict perceptual assimilation patterns which is in line with the findings of Ghaffarvand Makari (2018) in his study of Azerabaijani learners of English.

To discuss the predictions for L2 vowel contrasts, there were cases of SUBSET, SIMILAR and NEW regarding both acoustic and perceptual assimilations. A SUBSET scenario means each member of a contrast is categorized as a different L1 category. There were cases of SUBSET scenario in both acoustic and perceptual assimilations for /ʊ-u/ and /ɑ-ʌ/ in which both members of any of these SSBE contrasts were categorized as more than a native category. In this contrast /ʊ/ was categorized as FA /u/ and /o/ most of the time and /u/ was categorized as either FA /u/ or /o/ with different percentages. In a similar fashion, the /ɑ-ʌ/ contrast was a case of SUBSET scenario in which both SSBE /ɑ/ and /ʌ/ were mostly assimilated to FA /ɒ/ and /æ/, respectively.

With regards to SIMILAR scenarios, this case was seen for both acoustic and perceptual assimilations for 3 SSBE contrasts: /i-ɛ/, /ɔ-u/ and /æ-ɑ/. In cases of SIMILAR scenarios, each member of an SSBE contrast is assimilated to a different L1 category. In SSBE /i-ɛ/ contrast, /i/ was mostly categorized as FA /i/ and /ɛ/ was categorized as FA /e/. In SSBE /ɔ-u/, /ɔ/ was perceptually assimilated to FA /o/ and acoustically to FA /u/, while SSBE /u/ was perceptually assimilated to FA /u/ and acoustically assimilated to FA /i/. In SSBE /æ-ɑ/ contrast, the SSBE /æ/ was mostly categorized as FA /æ/ both perceptually and acoustically whilst SSBE /ɑ/ was acoustically and perceptually assimilated to FA /ɒ/.

Acoustic and perceptual assimilations yielded different results for these contrasts: /ɔ-ɒ/, /æ-ʌ/, /æ-ɜ/ and /i-ɪ/. The /ɔ-ɒ/ contrast was a case of NEW scenario when perceptually tested but was a case of SIMILAR scenario when tested acoustically. The /æ-ʌ/ contrast was a case of SUBSET perceptually but it turned out to be a case of SIMILAR scenario acoustically. Similarly, the /æ-ɜ/ was a case of SUBSET scenario perceptually and a case of SIMILAR scenario acoustically. The /i-ɪ/ contrast was a case of SIMILAR scenario perceptually and a case of NEW scenario acoustically.

Finally, it is valuable to briefly refer to the discrimination results once again. That section provided us with the following information: the contrast that proved to be the easiest to discriminate was /æ-ɑ/. Other contrasts which were easy to discriminate were /ɔ-u/, /i-ɛ/, /i-ɪ/. Those contrasts which were moderately easy to discriminate were /æ-ɜ/, /æ-ʌ/, /ɔ-ɒ/ and /ɑ-ʌ/.

Lastly, the most difficult contrast to discriminate was /ʊ-u/.

In this study the acoustic cues, effective in discriminations of our participants, were not investigated. In addition to that, this study did not focus on participants' reliance of vowel duration or spectral difference during perception of L2 vowel pairs. This calls for more thorough investigation of those factors that have an effect on L2 speech perception, suggesting that in future studies reliance of spectral and duration could be further investigated.

8 Conclusion

At the start of this project I found no conclusive and structured study about how difficult it is for Farsi learners of English when perceiving English vowels. As a result, this study has aimed to concentrate on this issue. We hope this study can be beneficial to teachers of English in Iran or everyone interested in this topic by providing them with a better insight on how Iranians with limited exposure to English as L2 would perceive standard southern British English vowels. Based on our results, more focus needs to be placed on teaching tense /lax vowels.

In this study, I have tried to predict discriminations of Standard Southern British English vowels by speakers of Farsi language who were naïve learners of English. A perceptual assimilation test was conducted and compared with the predictions based on acoustic properties. These comparisons show that our assimilation results are to some extent in line with those of acoustic results, although there were some discrepancies at times.

The null hypothesis was rejected as Farsi speakers were indeed able to distinguish L2 vowels at varying degrees. Research questions 1 to 3 were answered by using assimilation and discrimination tasks and were discussed in details in designated sections. Research question 4 was answered by comparing acoustic data to those of assimilation and discrimination tasks.

Finally, it seems possible to collect both acoustic data and assimilation data from the same group of listeners in a similar fashion in order for the L2LP model to lead into more precise and accurate results.

Additionally, this study could further be extended by examining the probable success of participants who undergo intensive phonetic training to see if this variable can relate to their overall musical abilities (Ghaffarvand Mokari, 2018). This will subsequently reveal if training will improve discrimination and production of L2 vowels. Another interesting topic would be having two languages side by side to see which one is better for discrimination of a target language. For example, Azerbaijani and Farsi language could be explored further by taking into consideration the vowel inventory size of each language in comparison to Target language and acoustic similarity of each language with target language and then observing the performance of participants in discrimination of target language contrasts to see which factor has more power of discriminating vowels of a target language such as English.

.

APPENDIX A

Information consent form

UNIVERSITY OF EASTERN FINLAND

پرسشنامه مطالعاتی

مشخصات عمومی

مطالعه تشابه ادراکی واکه هاي انگلیسی بریتانیایی توسط بومی هاي فارسی

شماره پرسشنامه:

تار یخ: / /

نیازي به درج نام و نام خانوادگی تان وجود ندارد .

لطفا به سوالات زیر به دقت پاسخ دهید. صحت اطلاعات شما ما را در انجام این تحقیق یاري می کند.

- جنسیت: مونث ☐ مذکر ☐
- تاریخ تولد: / /
- شهر محل تولد: ....
- تحصیلات: ....
- شغل: ....: ....
- زبان مادري: ....
- گویش: تهرانی ☐ اگر گویش شما غیر تهرانی است لطفا ذکر کنید... :
- زبان مادري پدرتان..........: ....
- زبان مادري مادرتان: ....
- آیا به زبان دیگري به غیر از زبان مادریتان تسلط دارید؟ لطفا نام زبان و میزان آشنایی خود را با نمرات 1 تا 7 ارزیابی کنید . (نمره1 به عنوان آشنایی در حد چند کلمه و 7 به عنوان تسلط در حد زبان مادري در نظر گرفته می شود.)

زبان غیر از زبان مادري که با آن آشنایی دارید: نام زبان: .... میزان آشنایی: 1 – 2 – 3 – 4 – 5 – 6 - 7

زبان غیر از زبان مادري که با آن آشنایی دارید: نام زبان: .... میزان آشنایی: 1 – 2 – 3 – 4 – 5 – 6 - 7

- آیا تا بحال در شهري بجز شهر کنونی خود به مدت بیش از شش ماه سکونت داشته اید؟ بلی ☐ کدام شهر:............. خیر ☐
- آیا سابقه اختلال شنوایی یا گفتاري داشته اید؟

با نهایت تشکر از همکاري شما

APPENDIX B

Information consent form

UNIVERSITY OF
EASTERN FINLAND

شماره ....

فرم رضایتنامه ي کتبی

بدینوسیله به اطلاع میرساند اینجانب .... پس از آگاهی کامل از هدف و شیوه هاي اجراي مطالعه توسط مجري طرح، رضایت کتبی خود را مبنی بر شرکت در طرح تحقیقاتی با عنوان مطالعه « تشابه ادراکی واکه هاي انگلیسی بریتانیایی توسط بومی هاي فارسی» اعلام میدارم .

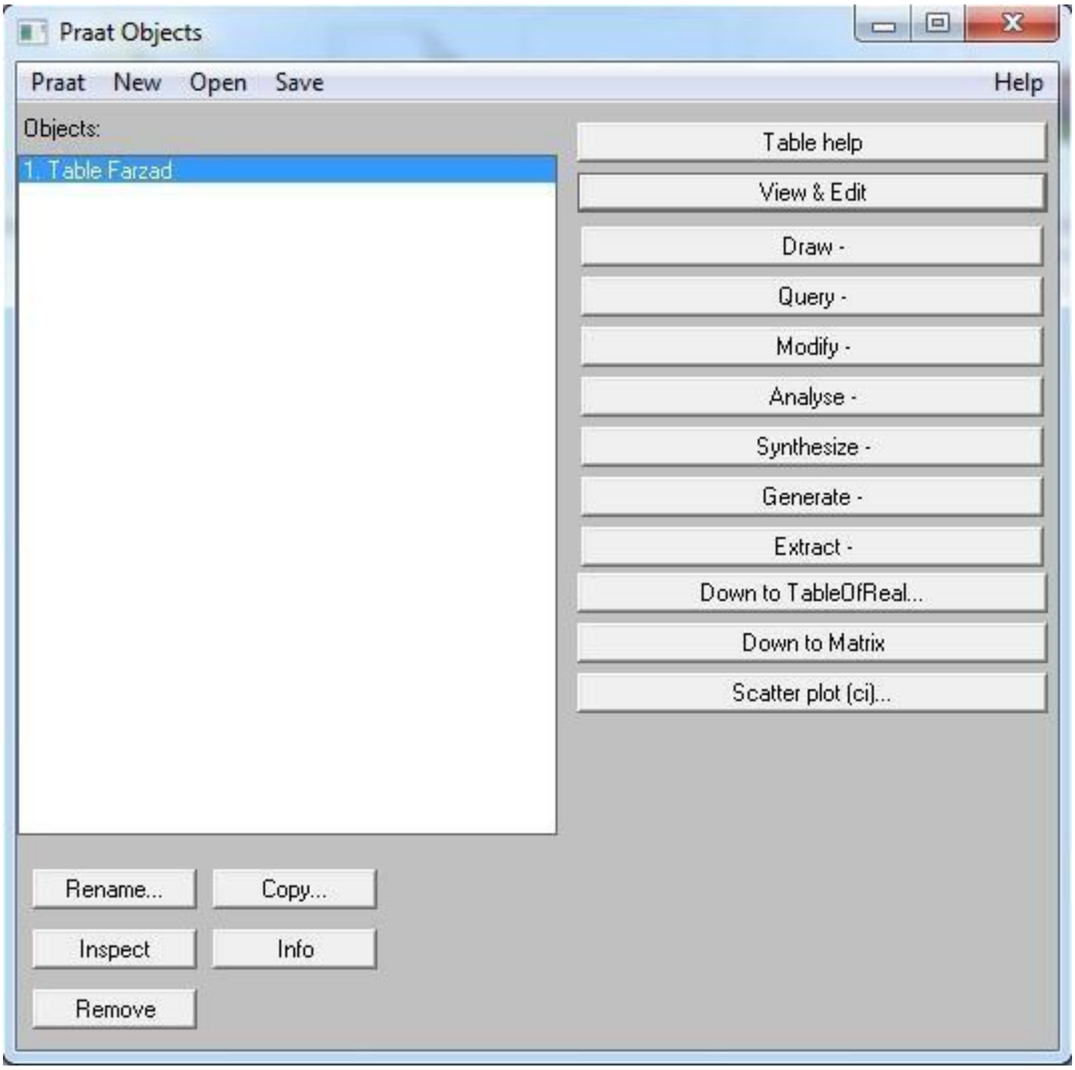
نام و نام خانوادگی..............................

تاریخ..

امضا و یا اثر انگشت

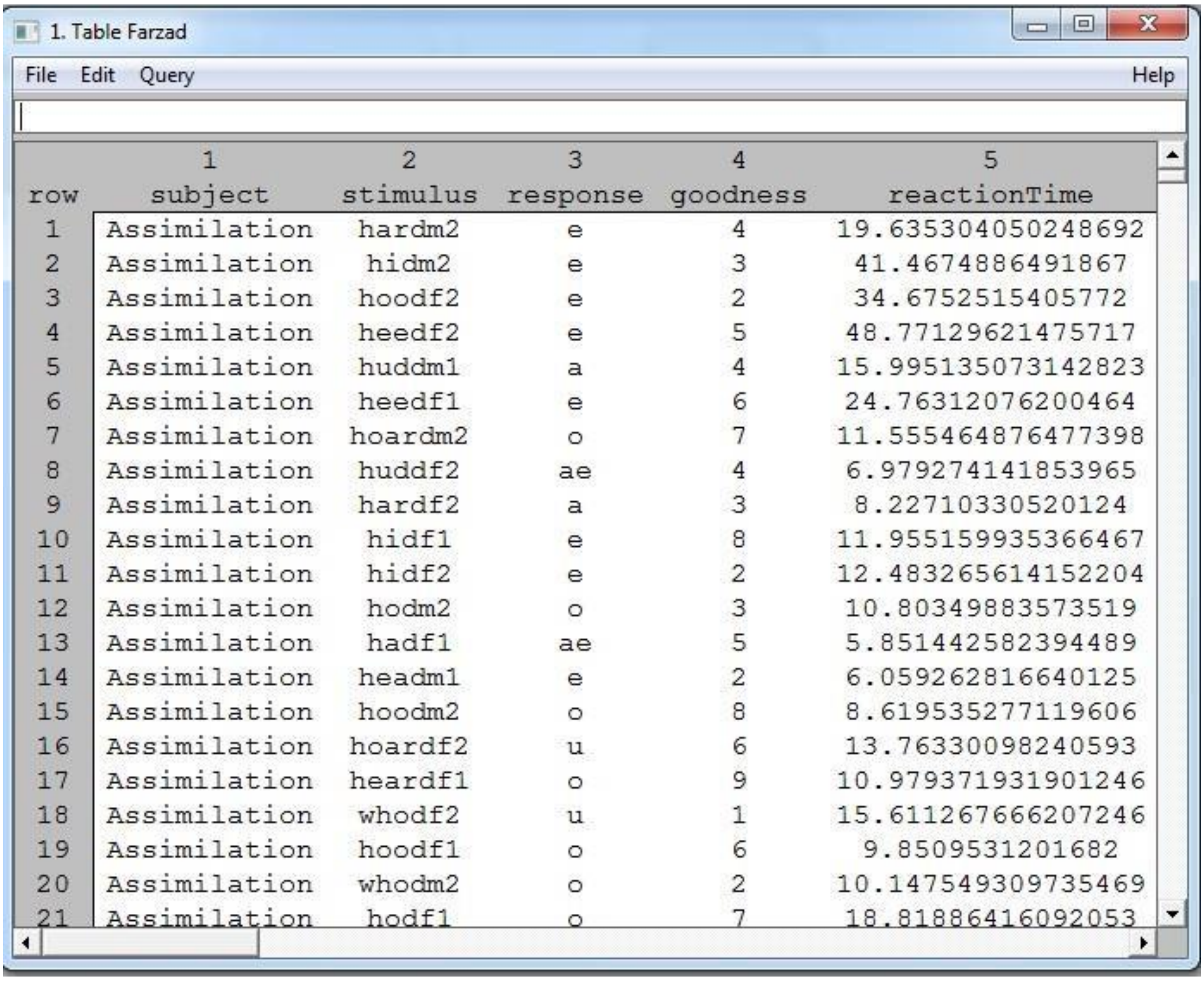
APPENDIX C

Screenshot of Praat



APPENDIX D

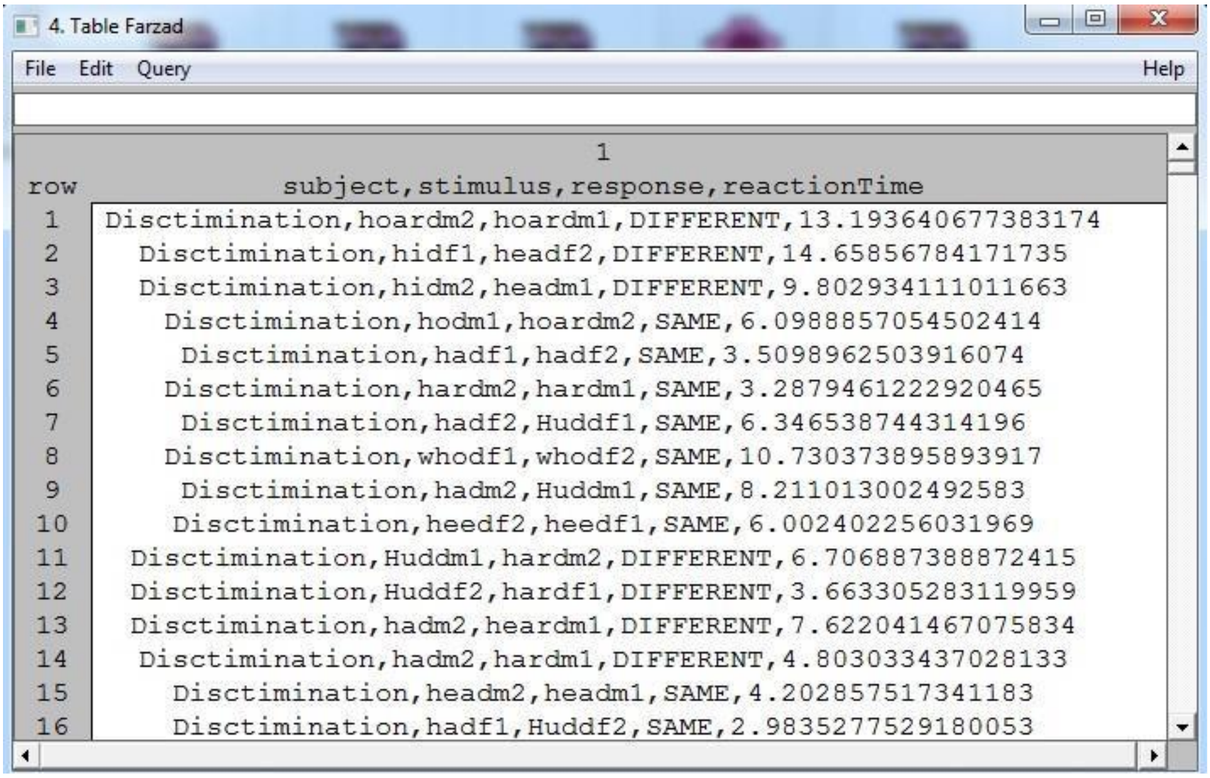
Screenshot of Praat



row	1 subject	2 stimulus	3 response	4 goodness	5 reactionTime
1	Assimilation	hardm2	e	4	19.635304050248692
2	Assimilation	hidm2	e	3	41.4674886491867
3	Assimilation	hoodf2	e	2	34.6752515405772
4	Assimilation	heedf2	e	5	48.77129621475717
5	Assimilation	huddm1	a	4	15.995135073142823
6	Assimilation	heedf1	e	6	24.76312076200464
7	Assimilation	hoardm2	o	7	11.555464876477398
8	Assimilation	huddf2	ae	4	6.979274141853965
9	Assimilation	hardf2	a	3	8.22710330520124
10	Assimilation	hidf1	e	8	11.955159935366467
11	Assimilation	hidf2	e	2	12.483265614152204
12	Assimilation	hodm2	o	3	10.80349883573519
13	Assimilation	hadf1	ae	5	5.851442582394489
14	Assimilation	headm1	e	2	6.059262816640125
15	Assimilation	hoodm2	o	8	8.619535277119606
16	Assimilation	hoardf2	u	6	13.76330098240593
17	Assimilation	heardf1	o	9	10.979371931901246
18	Assimilation	whodf2	u	1	15.611267666207246
19	Assimilation	hoodf1	o	6	9.8509531201682
20	Assimilation	whodm2	o	2	10.147549309735469
21	Assimilation	hodf1	o	7	18.81886416092053

APPENDIX E

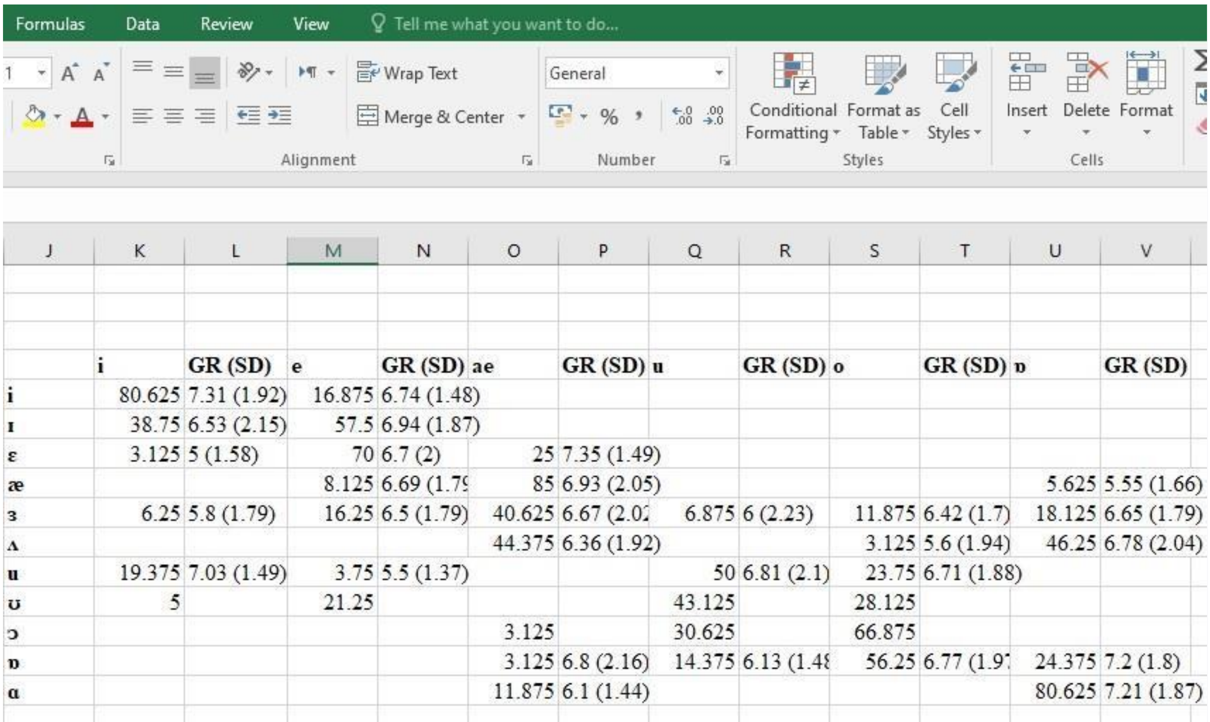
Screenshot of Praat



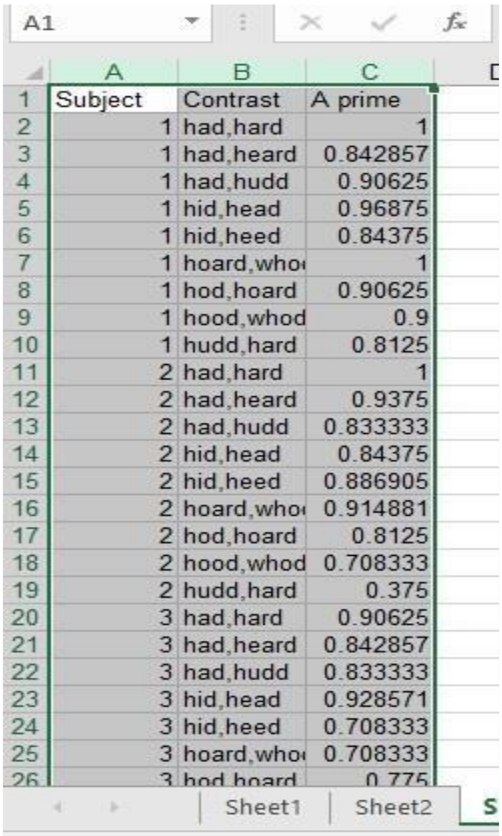
row	1 subject,stimulus,response,reactionTime
1	Disctimination,hoardm2,hoardm1,DIFFERENT,13.193640677383174
2	Disctimination,hidf1,headf2,DIFFERENT,14.65856784171735
3	Disctimination,hidm2,headm1,DIFFERENT,9.802934111011663
4	Disctimination,hodm1,hoardm2,SAME,6.0988857054502414
5	Disctimination,hadf1,hadf2,SAME,3.5098962503916074
6	Disctimination,hardm2,hardm1,SAME,3.2879461222920465
7	Disctimination,hadf2,Huddf1,SAME,6.346538744314196
8	Disctimination,whodf1,whodf2,SAME,10.730373895893917
9	Disctimination,hadm2,Huddm1,SAME,8.211013002492583
10	Disctimination,heedf2,heedf1,SAME,6.002402256031969
11	Disctimination,Huddm1,hardm2,DIFFERENT,6.706887388872415
12	Disctimination,Huddf2,hardf1,DIFFERENT,3.663305283119959
13	Disctimination,hadm2,heardm1,DIFFERENT,7.622041467075834
14	Disctimination,hadm2,hardm1,DIFFERENT,4.803033437028133
15	Disctimination,headm2,headm1,SAME,4.202857517341183
16	Disctimination,hadf1,Huddf2,SAME,2.9835277529180053

APPENDIX F

Screenshot of Excel



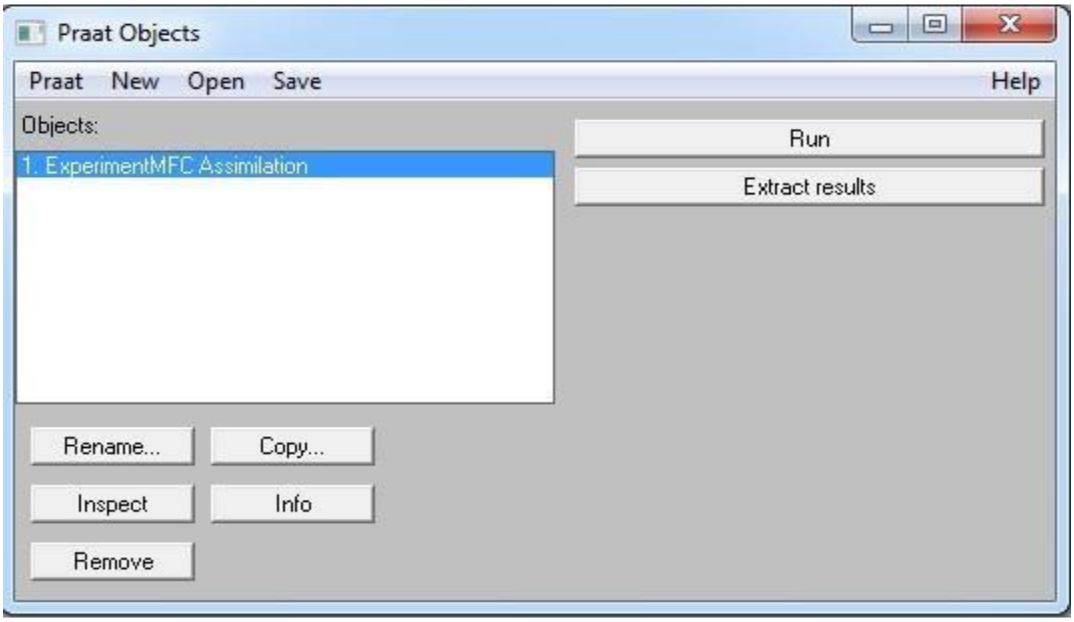
J	K	L	M	N	O	P	Q	R	S	T	U	V
	i	GR (SD)	e	GR (SD)	ae	GR (SD)	u	GR (SD)	o	GR (SD)	ɒ	GR (SD)
i	80.625	7.31 (1.92)	16.875	6.74 (1.48)								
ɪ	38.75	6.53 (2.15)	57.5	6.94 (1.87)								
ɛ	3.125	5 (1.58)	70	6.7 (2)	25	7.35 (1.49)						
æ			8.125	6.69 (1.7	85	6.93 (2.05)					5.625	5.55 (1.66)
ɜ	6.25	5.8 (1.79)	16.25	6.5 (1.79)	40.625	6.67 (2.0	6.875	6 (2.23)	11.875	6.42 (1.7)	18.125	6.65 (1.79)
ʌ					44.375	6.36 (1.92)			3.125	5.6 (1.94)	46.25	6.78 (2.04)
u	19.375	7.03 (1.49)	3.75	5.5 (1.37)			50	6.81 (2.1)	23.75	6.71 (1.88)		
ʊ	5		21.25				43.125		28.125			
ɔ					3.125		30.625		66.875			
ɒ					3.125	6.8 (2.16)	14.375	6.13 (1.4	56.25	6.77 (1.9	24.375	7.2 (1.8)
ɑ					11.875	6.1 (1.44)					80.625	7.21 (1.87)

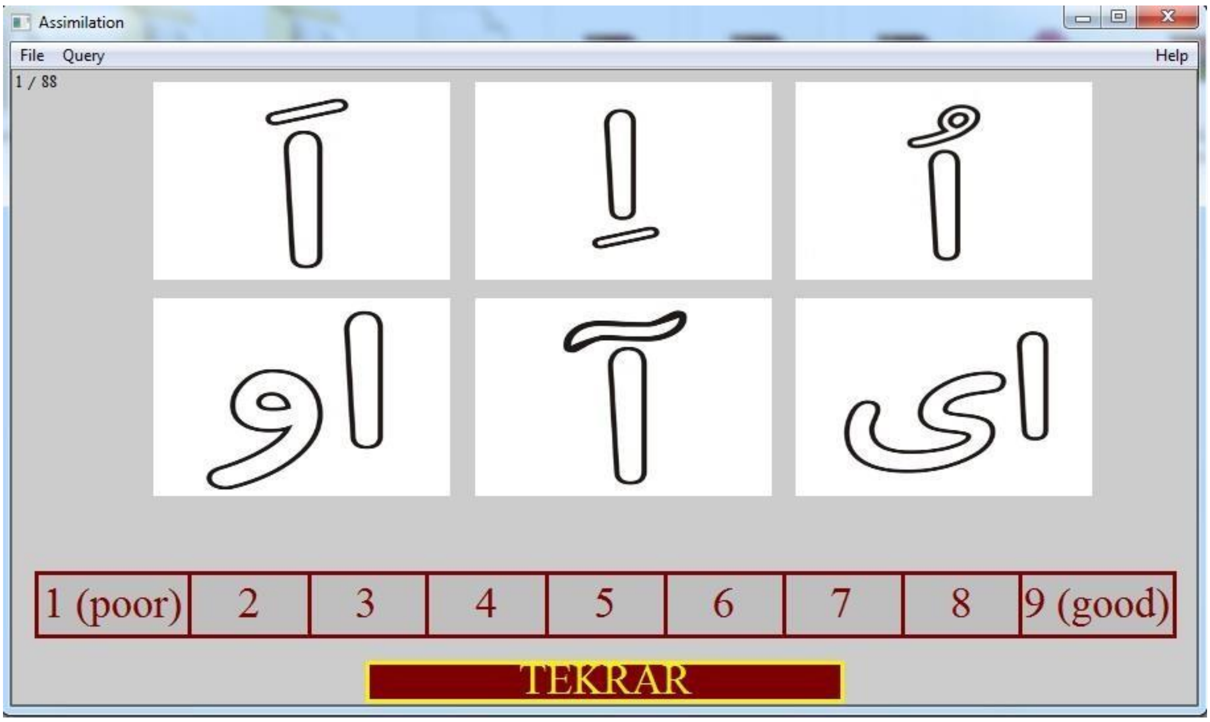


	A	B	C
1	Subject	Contrast	A prime
2	1	had,hard	1
3	1	had,heard	0.842857
4	1	had,hudd	0.90625
5	1	hid,head	0.96875
6	1	hid,heed	0.84375
7	1	hoard,who	1
8	1	hod,hoard	0.90625
9	1	hood,whod	0.9
10	1	hudd,hard	0.8125
11	2	had,hard	1
12	2	had,heard	0.9375
13	2	had,hudd	0.833333
14	2	hid,head	0.84375
15	2	hid,heed	0.886905
16	2	hoard,who	0.914881
17	2	hod,hoard	0.8125
18	2	hood,whod	0.708333
19	2	hudd,hard	0.375
20	3	had,hard	0.90625
21	3	had,heard	0.842857
22	3	had,hudd	0.833333
23	3	hid,head	0.928571
24	3	hid,heed	0.708333
25	3	hoard,who	0.708333
26	3	hod,hoard	0.775

APPENDIX G

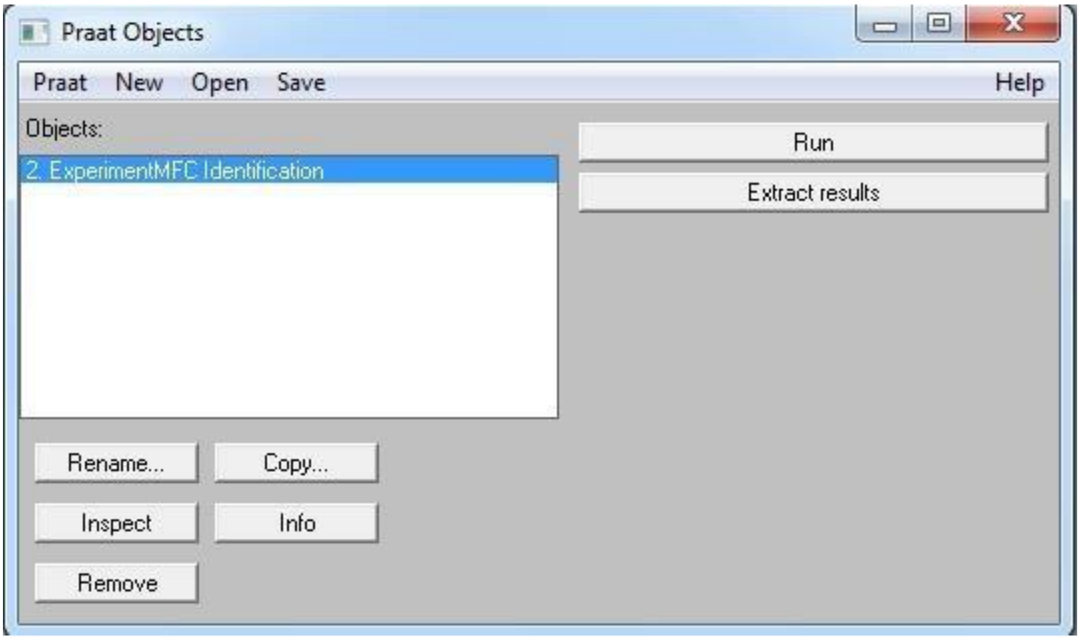
Screenshot of Praat

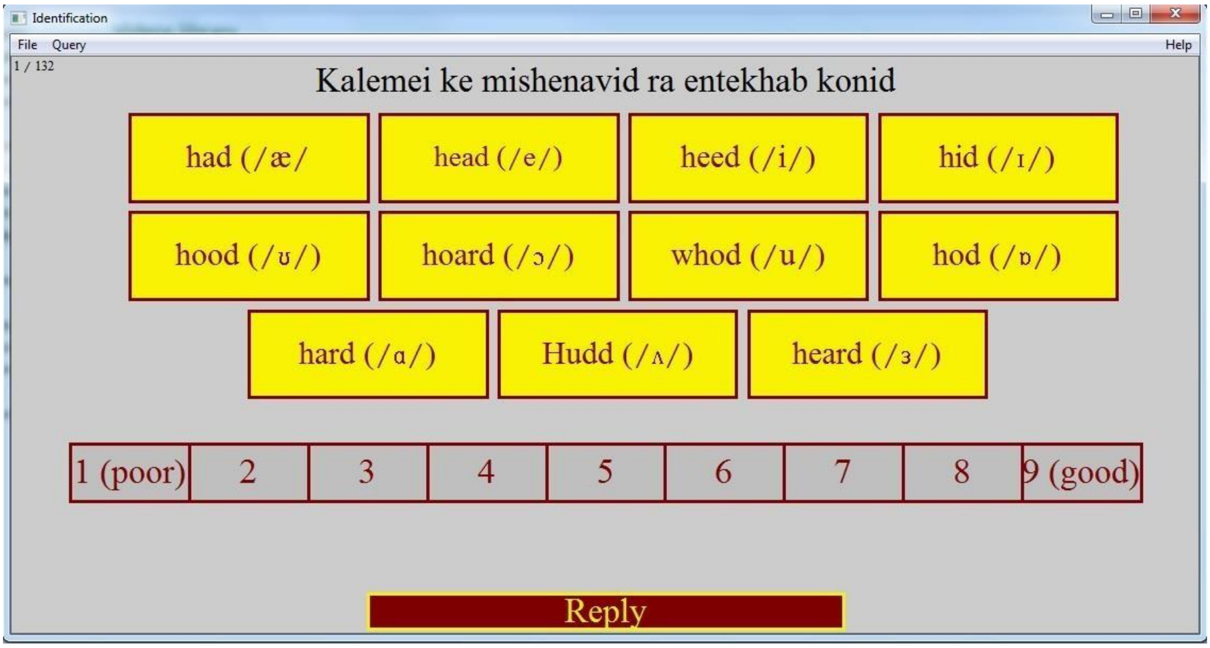




APPENDIX H

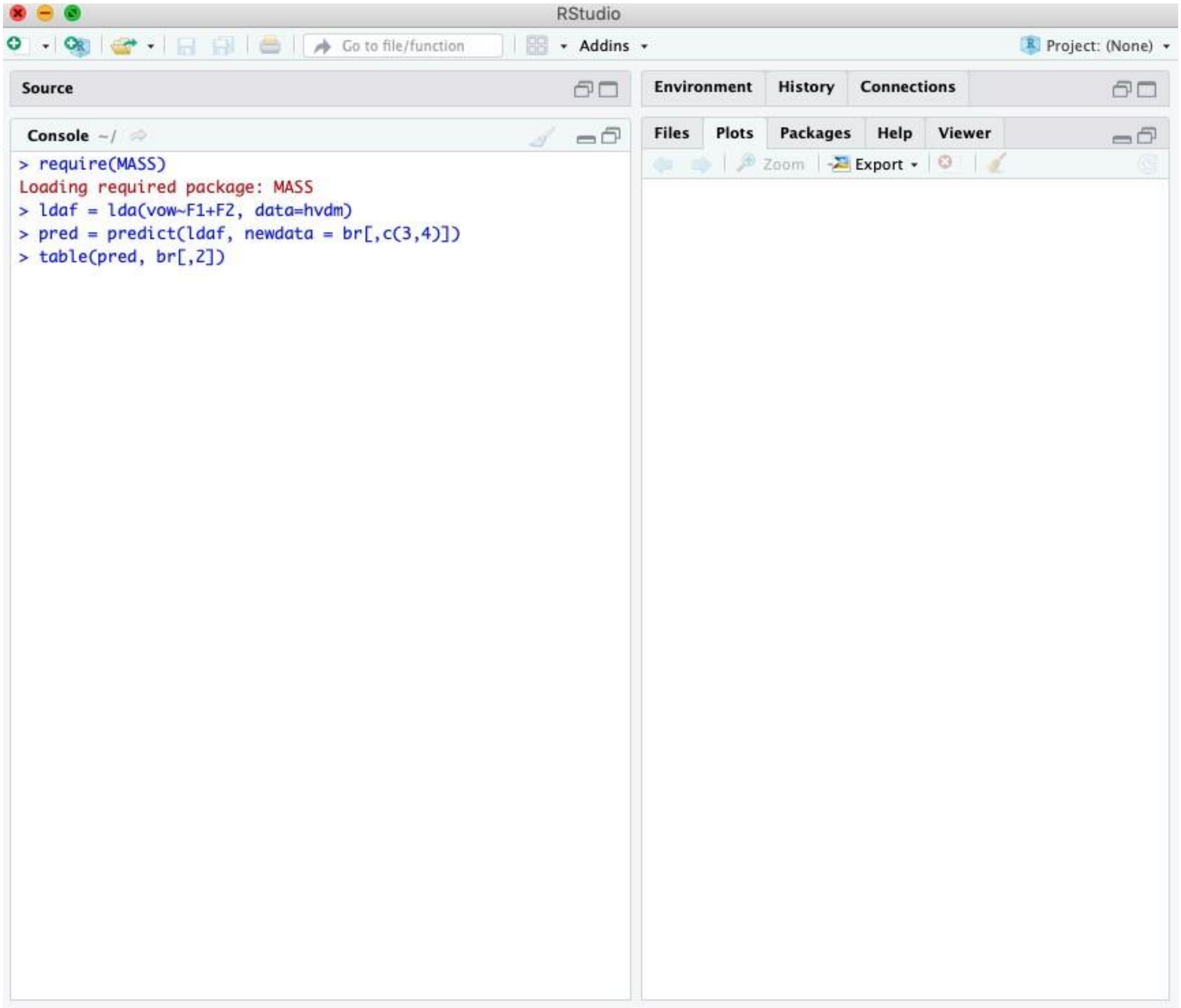
Screenshot of Praat





APPENDIX I

A SCREENSHOT OF RSTUDIO SOFTWARE USED FOR ANALYSING CROSS-LINGUISTIC LDA



hdvm refers to Farsi vowel formants and *br* refers to British vowel formants

Bibliography

Ansarin, Ali Akbar. "An acoustic analysis of modern Persian vowels." *9th Conference Speech and Computer*. 2004.

Best, Catherine T. "A direct realist view of cross-language speech perception." *Speech perception and linguistic experience* (1995): 171-204.

Best, Catherine T. "The emergence of native-language phonological influences in infants: A perceptual assimilation model." *The development of speech perception: The transition from speech sounds to spoken words* 167.224 (1994): 233-277.

Best, Catherine T. "Learning to perceive the sound pattern of English." *Advances in infancy research* 9 (1995): 217-217.

Best, Catherine T., Gerald W. McRoberts, and Elizabeth Goodell. "Discrimination of nonnative consonant contrasts varying in perceptual assimilation to the listener's native phonological system." *The Journal of the Acoustical Society of America* 109.2 (2001): 775-794.

Best, Catherine T., and Michael D. Tyler. "Nonnative and second-language speech perception: Commonalities and complementarities." *Language experience in second language speech learning: In honor of James Emil Flege* 1334 (2007): 13-34.

Bion, Ricardo AH, et al. "Category formation and the role of spectral quality in the perception and production of English front vowels." *Ninth International Conference on Spoken Language Processing* (2006): 1363-1366.

Boersma, Paul, and D. Weenink. "Praat: Doing phonetics by computer [version 6.0. 06]." (2015).

Cebrian, Juli. "Experience and the use of duration in the categorization of L2 vowels." *Journal of Phonetics* 34 (2006): 372-387.

Cebrian, Juli. "The effect of perceptual factors in the acquisition of an L2 vowel contrast." *Contrast in Phonology: Theory, Perception, Acquisition* 13 (2008): 303-321.

Cebrian, Juli, Joan C. Mora, and Cristina Aliaga-Garcia. "Assessing crosslinguistic similarity by means of rated discrimination and perceptual assimilation tasks." *New Sounds 2010: Proceedings of the Sixth International Symposium on the Acquisition of Second Language Speech (*2010): 77-82

Chládková, Kateřina, and Václav Jonáš Podlipský. "Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners." *The Journal of the Acoustical Society of America* 130.4 (2011): EL186-EL192.

Colantoni, Laura, et al. *Second language speech*. Cambridge University Press, 2015.

Derwing, Tracey M., and Murray J. Munro. *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Vol. 42. John Benjamins Publishing Company, 2015.

Deterding, David. "The North Wind versus a Wolf: short texts for the description and measurement of English pronunciation." *Journal of the International Phonetic Association* (2006): 187-196

Elvin, Jaydene, Paola Escudero, and Polina Vasiliev. "Spanish is better than English for discriminating Portuguese vowels: acoustic similarity versus vowel inventory size." *Frontiers in psychology* 5 (2014): 1188.

Escudero, Paola. *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Netherlands Graduate School of Linguistics, 2005.

Escudero, Paola. "Linguistic perception of “similar” L2 sounds." *Phonology in perception* 15 (2009): 152-190.

Escudero, Paola, and Paul Boersma. "Bridging the gap between L2 speech perception research and phonological theory." *Studies in Second Language Acquisition* (2004): 551-585.

Escudero, Paola, and Kateřina Chládková. "Spanish listeners’ perception of American and Southern British English vowels. " *The Journal of the Acoustical Society of America* 128.5 (2010): EL254-EL260

Fabra, Lucrecia Rallo. *Acquisition of a second language vowel system: the case of catalan learners of english*. Diss. Universitat de Barcelona, 2005.

Fabra, Lucrecia Rallo, and Joaquín Romero. "Native Catalan learners' perception and production of English vowels." *Journal of Phonetics* 40.3 (2012): 491-508.

Flege, James E. "Second language speech learning: Theory, findings, and problems." *Speech perception and linguistic experience: Issues in cross-language research* 92 (1995): 233-277.

Flege, James E. "Methods for assessing the perception of vowels in a second language." *Issues in clinical linguistics* (2003): 19-44.

Flege, James Emil, Naoyuki Takagi, and Virginia Mann. "Lexical familiarity and Englishlanguage experience affect Japanese adults' perception of/ɹ/and/l." *The Journal of the Acoustical Society of America* 99.2 (1996): 1161-1173.

Flege, James Emil, Ocke-Schwen Bohn, and Sunyoung Jang. "Effects of experience on nonnative speakers' production and perception of English vowels." *Journal of phonetics* 25.4 (1997): 437-470.

Ghaffarvand Mokari, P (2018). Second language vowel learning and the role of cognitive abilities" (unpublished doctoral dissertation). University of Eastern Finland.

Gilichinskaya, Yana D., Franzo Law, and Winifred Strange. "Speeded discrimination of American vowels by experienced Russian learners of English." *The Journal of the Acoustical Society of America* 122.5 (2007): 3028.

Gilichinskaya, Yana D., and Winifred Strange. "Perceptual assimilation of American English vowels by inexperienced Russian listeners." *The Journal of the Acoustical Society of America* 128.2 (2010): EL80-EL85.

Goto, Hiromu. "Auditory perception by normal Japanese adults of the sounds" L" and" R."." *Neuropsychologia* (1971): 317-323.

Hawkins, Sarah, and Jonathan Midgley. "Formant frequencies of RP monophthongs in four age groups of speakers." *Journal of the International Phonetic Association* 35.2 (2005): 183-199.

Hudson, Grover. *Essential introductory linguistics*. Wiley-Blackwell, 1999.

Hunter, Gordon, and Hanna Kebede. "Formant frequencies of British English vowels produced by native speakers of Farsi." *Acoustics 2012*. 2012.

Iverson, Paul, Valerie Hazan, and Kerry Bannister. "Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English/r/-/l/to Japanese adults." *The Journal of the Acoustical Society of America* 118.5 (2005): 3267-3278.

Jeske, Andrew. *The perception of English vowels by native Spanish speakers*. Diss. University of Pittsburgh, 2012.

Ladefoged, Peter, and Keith Johnson. *A course in phonetics*. Cengage learning, 2014, pp. 21-23.

Lengeris, Angelos. "Perceptual assimilation and L2 learning: Evidence from the perception of Southern British English vowels by native speakers of Greek and Japanese." *Phonetica* 66.3 (2009): 169-187.

Levy, Erika S. "On the assimilation-discrimination relationship in American English adults' French vowel learning." *The Journal of the Acoustical Society of America* 126.5 (2009): 2670-2682.

Levy, Erika S. "Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French." *The Journal of the Acoustical Society of America* 125.2 (2009): 1138-1152.

McAllister, Robert, James E. Flege, and Thorsten Piske. "The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian." *Journal of phonetics* 30.2 (2002): 229-258.

Mora, Joan Carles. "Ocke-Schwen Bohn and Murray J. Munro (eds)(2007): Language Experience in Second Language Speech Learning: In Honor of James Emile Flege, Language learning and language teaching monograph series, volume 17, Amsterdam, John Benjamins." *Estudios de fonética experimental* (2007): 287-298.

Mokari, Payam Ghaffarvand, Stefan Werner, and Ali Talebi. "An acoustic description of Farsi vowels produced by native speakers of Tehrani dialect." *The Phonetican Journal of the International Society of Phonetic Sciences* 114 (2017): 6-23.

Nishi, Kanae, et al. “Acoustic and perceptual similarity of Japanese and American English vowels.” The journal of the Acoustic Society of America 124.1 (2008): 576-588.

Pereira Reyes, Y. I. *Perception and production of English vowels by Chilean learners of English: effect of auditory and visual modalities on phonetic training*. Diss. UCL (University College London), 2014.

Roach, Peter. "British English: Received pronunciation." *Journal of the International Phonetic Association* 34.2 (2004): 239-245.

Snodgrass, Joan Gay, Gail Levy-Berger, and Martin Haydon. *Human experimental psychology*. Vol. 395. New York: Oxford University Press, 1985.

Strange, Winifred, et al. "Acoustic and perceptual similarity of North German and American English vowels." *The Journal of the Acoustical Society of America* 115.4 (2004): 1791-1807.

Wang, Hongyan, and Vincent J. Van Heuven. "Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers." *Linguistics in the Netherlands* 23.1 (2006): 237-248.

Yule, George. "The Study of Language, Cambridge University Press." *NY: USA* (2006).