**Review of dissertation**

*Production and Perceptual Representation of American English Vowel Sounds by Monolingual Persian and Early Bilingual Azerbaijani-Persian Adolescents*

Naeimeh Afshar

Supervisor: Prof. Dr. Vincent J. van Heuven

Reviewer: Szilárd Szentgyörgyi

Naeimeh Afshar's PhD dissertation is a well-researched, well-designed, logically structured work based on and carefully carried out research and illustrated and supported by several sets of experimental data and a detailed survey of the relevant literature. already at the beginning of the review, I have to say that I consider the dissertation very valuable and I suggest that the candidate should be awarded the PhD degree after a successful defence of her work.

First, I am going to describe the structure of the dissertation to be followed by my comments concerning the main points that I have found not completely satisfactory or where my professional opinion is different from the one reflected by the dissertation.

The dissertation is 127 pages long and complemented by 15 appendices throughout almost 30 more pages. It has been structured into 7 major chapters that discuss the following topics:

- **Chapter 1** contains the general introduction and a discussion of the state of affairs concerning languages spoken in Iran, as well as an introduction to bilingualism, he nature of the acquisition of non-native sounds, and the importance of vowels in pronunciation.
- **Chapter 2** is a thorough review of the relevant literature and serves as a theoretical background to the rest of the dissertation. It includes the topics of a comparative introduction to the vowel systems of the three languages studied: American English, Persian, and Azerbaijani, the acquisition of third language phonology – in comparison to that of second language phonology –, the relevance of perceptual vowel studies for foreign language learning/teaching, the relationship between the perception and production of L2 speech sounds, as well as language dominance. It is in this section that the research questions and hypotheses are introduced in detail.
- **Chapter 3** is completely dedicated to the discussion of language dominance in Persian-Azerbaijani bilinguals and the application and results of the LEAP-Q test to establish language dominance relations in the participating bilingual research subjects.

- **Chapter 4** is a perceptual assimilation study determine how the monophthongs of American English are perceptually assimilated by– both monolingual and bilingual – EFL learners in Iran. the author strives to find out and prove whether the monolingual Persian EFL learners assimilate the American English monophthongs according to the same patterns as bilingual EFL learners do when the latter are instructed to pair the American English vowels with those of Persian. Finally, the author wishes to determine if vowels of English assimilate in the same or in a different way to vowels shared between Persian and Azerbaijani as these results could indicate if the vowels only found in Azerbaijani affect the task performance in the test positively or negatively.
- **Chapter 5** discusses the mapping of the perceptual vowel spaces in the participants' native and foreign languages to point out the mental conceptions monolingual Persian and bilingual Azerbaijani/Persian learners of English have of the American English vowel system in terms of the vowel quality (color) and vowel duration compared with that of native speakers of American English.
- **Chapter 6** presents a contrastive acoustic analysis of vowels since it is critical that such cross-linguistic perceptual similarities be established in order to predict L2 learning difficulties more accurately. For this reason, vowel duration, the F1 and F2 formants representing vowel height and backness/rounding respectively, were measured in the recordings made of the participants' pronunciation of American English vowels in carrier words, and then compared with similar data collected from American L1 speakers.
- **Chapter 7**, the last section of the dissertation, presents a discussion of the most important results from the preceding chapters and draws the conclusions compared to the hypotheses discussed in Chapter 2.

I have to say that the methodology used and detailed conclusions are present in all chapters mostly without any major problematic issues proving that the author has invested the appropriate amount of work into her research and the output quality clearly meets the standard requirements of the Multilingualism Doctoral School.

Let me now present my notes and remarks for each of the sections where there was something either missing or not dealt with carefully enough in my opinion. For each section, I will first

explain my general theoretical or methodological remarks, then point out any problematic detail or minor shortcoming I may have found.

**Chapter 1**

There are several issues connected to the set of vowels introduced in this chapter as the set to be investigated in the dissertation as well as with the way some of these vowels are being characterized.

First of all, the dissertation aims at examining how the monophthongs of American English are perceived and pronounced by monolingual Persian and bilingual Persian-Azerbaijani speakers. There are two major areas in this characterization of the objectives of the dissertation: (1) there is no such thing as American English, and (2) some of the vowels listed as monophthongs in the dissertation are realized as diphthongs and are normally not pronounced as monophthongs in American English accents.

When the author refers to American English or General American English, it should be emphasized exactly which American English she is referring to precisely. From the name used as General American English, I take that it is meant to refer to General American, the variety often used to characterize the general Midwest accent spoken in Iowa or Missouri, for instance but it is contradicted by the fact that on page 60 in Chapter 5, the author claims that "American native control listeners (*N* = 20), all of whom spoke a form of General American English. These native speakers hailed from many different states in the USA, although half of them were born and bred in California". This, however, surely means that the American participants did speak (somewhat) different varieties of American English and thus, one may expect different vowel qualities for the same vowel phonemes from them and they may also be expected to perceive the same input in different ways, which is quite a methodological problem.

On the other hand, the vowels listed as the focus of attention, i.e. /i, e, æ, ɑ, ɔ, o, u, ɪ, ɛ, ʌ, ʊ/ are often mischaracterized: two of these vowels, /e/, or rather /eɪ/, and /o/, or rather /oʊ/, are actually pronounced as diphthongs in General American and are not subject to monophthongization in this dialect – and they are not monophthongized in most North American dialects either. Monophthongization of these vowels is very common in some dialects – e.g. northern varieties spoken in the U.K., e.g. Scottish English. For this reason, I think it is a mistake to treat it as a monophthong regardless of the longstanding tradition among some American linguists to do so. Note that many other monophthong vowels of General American in this list are also often diphthongized: /iː, uː/ are very often pronounced as /ij, uw/ respectively as some alternative transcriptions used for General American often show. Also, it is not /eɪ, oʊ/

that are most often pronounced as monophthongs in American Englishes but rather the wide diphthongs /aɪ, aʊ/ as in Coastal Southern and Southern Mountain words like *time* /taɪm/ and *bounty* /ˈbaʊnti/ are pronounced as [taːm] and [ˈbaːn̄i] respectively. It seems inconsistent to include /eɪ, oʊ/ in the set of monophthongs even though they are not pronounced as such in this variety while those vowels that may be occasionally monophthongized – i.e. /aɪ, aʊ/ – are uniformly treated as diphthongs. In later chapters, the author does mention that some of the results seem to reflect the fact that /eɪ, oʊ/ seem to be diphthongs.

The last general note concerning the description of vowels in Chapter 1 – and also in later chapters – is the use of the tense-lax dichotomy in particular and the uncertainty concerning whether certain parts of the text are concerned with the mental/phonemic/underlying representations – what you would traditionally call phonemes – or with the actual reality of speech sounds, i.e. the physical/phonetic/surface representation – what you would call phones. Should the author use such a controversial pair of terms as tense vs lax, it would have been wise to include a longer introduction to these concepts and their use in a phonological vs a phonetic sense. Since it is not done anywhere in the text in detail, the reader is left alone with this question. Tenseness-laxness in a phonetic sense is simply a matter of more articulatory muscle tension vs the lack of it resulting in more extreme articulations of vowels closer to the edges of the vowel space in phonetically tense vowels vs more centralized articulations in phonetically lax ones. On the other hand, phonological tenseness and laxness are concerned with the positions the vowels occur in – e.g. tense vowels are allowed to appear in word final position, as in *may* /meɪ/, *two* /tuː/, *see* /siː/, while (stressed) lax vowels are never allowed to appear in such positions, hence no English words end with them, */præ/, */spɛ/, */kwɒ/, */plʌ/.

Another similar problem arises in connection with vowel length, an important issue as the perceptual classification of vowels by non-native speakers in the experiments show that they often judge vowels on the basis of their length and not their quality (colour). It is well-known that English vowels are classified into short vs long but that the length of so-called long vowels – i.e. that of phonologically/underlyingly long vowels – is unstable as they often shorten whenever followed by fortis – i.e. phonologically voiceless – consonants as in *beat* /biːt/-[bit] or *boot* /buːt/-[but]. Since vowel length is so unstable, it does matter quite a lot whether by "short vowel" we mean only phonemically short vowels like /ɪ, ʊ, ʌ, ɛ, æ, ɒ/ or we also include the shortened realizations of phonemically long vowels as the ones in the abovementioned sample words. Also, as I will point out later, it is this particular length alternation that could have shed some more light on how exactly length figures into the perception of English vowels by non-native speakers and that a great opportunity has been missed here for the lack of this

kind of examples and a comparison between how language learners react when the length difference as a cue is taken away as in pairs like *beat* /biːt/-[bit] vs *bit* /bɪt/-[bɪt]. As it is argued in paragraph 1 on page 10, "[s]pecifically, American native listeners rely much less on vowel duration as a correlate of the tense-lax distinction than the Iranian EFL learners do" – this is a necessary consequence of the fact that length difference is neutralized in positions before fortis consonants, native speakers can only rely on the colour difference between vowel pairs.

*Minor remarks:*

Section 1.8, page 7

- paragraph 2: "Lip rounding is unmarked (back=rounded, front=unrounded)" the back low vowel listed two lines higher is /ɑ/ instead of the rounded /ɒ/.
- paragraph3 line 6: "vowels in English are reduced to either schwa [ə] or [ɪ]". Actually, vowels may also be reduced to [ʊ]

**Chapter 2**

Chapter 2 continues to suffer from problems concerning the treatment of vowel length. The main question concerning this is whether it is phonetic or phonological length that is a more important cue for vowel identification: as I have already pointed out, vowel length is not a reliable predictor as long vowels regularly shorten in some positions. On the other hand, there is another kind of length problem left unaddressed here: the vowel /æ/ is treated as if it was a regular long vowel just like /ɔː, iː, uː/. However, it is clearly different in that it is not licensed in the same types of positions: /ɔː, iː, uː/ are allowed word finally while /æ/ is not. For this reason, it would be best to treat /æ/ as a phonologically short vowel which has a phonetically long allophone whenever it is not followed by fortis consonants.

The other issue that is brought up in this section is the basis for treating central or centralized vowels differently in English and Azerbaijani. In American English, the author claims that /ɪ, ɛ, ʌ, ʊ/ are lax because they are more central in articulation and, thus, there is a tense-lax distinction in American English. Why is the same kind of difference not claimed for Azerbaijani? Since the author does say that Azerbaijani also has central vowels – similarly to English –, it should logically follow that there is a tense-lax distinction phonetically. This can be treated by simply claiming that tense-lax is not a relevant distinction in English phonetically – it is a relevant phonological difference concerning the positions phonologically tense and lax vowels are allowed to occur in, e.g. in the case of vowel shift phenomena like *sane* [seɪn] – *sanity* [ˈsænəɾi], *holy* [ˈhoʊli] – *holiday* [ˈhɑlədeɪ] etc. If it is only a phonologically relevant

feature but not phonetically, then it should not be referred to when describing American English. Since the same – i.e. tense-lax is not a relevant phonetic distinction – is true in Azerbaijani, then it should not be used for that language either. This way, two languages displaying similar phonetic distinctions would be treated similarly.

*Minor remarks:*

- Page 14 Figure 2.1.B The symbol [œ] is used for the mid-low front rounded vowel, the mid-high front rounded vowel is mostly represented by [ø]

**Chapter 4**

Since I have no relevant remarks for chapter 3, we now turn to chapter 4. When discussing the differences between the three languages in Figure 4.1, it is mentioned that Azerbaijani and Persian do not have length and tenseness distinctions but it is claimed that American English does have both. However, as I have pointed out above Azerbaijani also has central/centralized vowels, thus, it is interesting that English and Azerbaijani vowels are not judged according to the same criteria – i.e. the fact that some vowels are more central(ized) than others.
Another problem appearing again is categorizing /e/ and /o/ as monophthongs, which I think is a conceptual mistake as they are not only diphthongized phonetically but they are underlying diphthongs that follow similar monophthongization patterns as other diphthongs in many varieties while there are no diphthongization patterns similar to the one claimed here for /e/ and /o/, with the possible exception of the Southern Drawl in words like dress /drɛss/-[dreɪəs]. The author claims that /e/ and /o/ only display cases of slight diphthongization not essential for their identification. But this may not be true in all varieties as in dialects where there is monophthongization of /eɪ/ and there is also raising of /ɛ/ to [e], /ɛ/ and /eɪ/ would be neutralized. This is also supported by the fact that even non-native speakers tend to identify /e/ as an /eɪ/ diphthong: as you note in paragraph 1 on page 54 "/ɛ/ is perceived as a much better exemplar (Fit-index: 3.6 = Fair) of Persian /e/ than either /ɪ/ (too high, Fit-index: 2.3 = Poor) or /e/ (too long and diphthongal, Fit-index: 2.1 = Poor) is". That is, Persian speakers tend to identify American English /e/ with Persian /e/ less as the former is rather identified as a diphthong, something that is missing from Persian.

**Chapter 5**

In this section, the question of which variety of American English the control listeners speak is brought up: „groups of American native control listeners ($N$ = 20), all of whom spoke a form

of General American English. These native speakers hailed from many different states in the USA, although half of them were born and bred in California". Since pronunciation of American Englishes may be quite different, differences in judgements by these speakers speaking different varieties must be expected. The question is how skewed is their judgement and whether we can draw any valid conclusions on the basis of this? That it would have been a good idea to pay more attention to the exact dialects that the American controls use is further supported by the generalization that it is not only Persian learners that fail to distinguish between the back vowels /ɑ, ɔ/ but also native speakers, whose representation is occupied by a merged category /ɑ, ɔ/. Thus, there must have been several different varieties of American English spoken by these controls, some displaying the *card-cod* neutralization only, while others also displaying the *caught-cot* merger.

Another methodological problem may arise as a result of the choice of the keywords used on the control panel of the sound identification task as shown in Figure 5.2: the keywords – *feel, fill, sale, tell, shall, null, fool, full, whole, call, doll* – had an /l/ following them. Using a dark /l/ to immediately follow the vowels is not a good idea since it is known to influence the pronunciation of the preceding vowel and make it produced further back than usual, which may distort the subjects' perception of these vowels, which could differ significantly from the perception of the same vowels before other consonants. This makes association of the vowels heard with the vowels in the keywords more difficult and less reliable.

## Chapter 6

No major concern has been found in this section. The only suggestion I would like to make is one concerning the last paragraph on page 84: "Incorrect pronunciations occurred in the vast majority of the tokens of *sawed* and *hawed*, which were then pronounced with a full diphthong /ɑʊ/ (as in *cloud*)." This phenomenon is very typical in language learners whose L1 does not have diphthongs and automatically think that diphthongs should be represented by vowel digraphs like <AW> or <AU>, and as a result, learners associate each individual vowel letter with a separate vowel quality. This, in turn, results in letter-by-letter pronunciation very common in the speech production of learners with other L1's – e.g. Hungarian learners of English often go through a phase of pronouncing words like because as [bɪˈkaʊz] instead of the correct pronunciation with a short vowel: [bɪˈkɒz] or [bɪˈkɔːz].

On the basis of the above discussion and remarks, I conclude that Naeimeh Afshar's PhD dissertation is valuable scientific work meeting all the requirements concerning form and content for doctoral dissertations in the Multilingualism Doctoral School in Linguistics at the Faculty of Modern Philology and Social Sciences at the University of Pannonia and I suggest that the author should be awarded the PhD degree if his defence procedure is successful in all respects.

October 21st, 2022, Veszprém

Szilárd Szentgyörgyi
University of Pannonia