

Answers to Dr. Szilárd Szentgyörgyi for his review on The perception and production of American English Sounds by Palestinian Arabic Adolescents”

I thank Prof. Szilárd Szentgyörgyi for all the very thorough and constructive comments he provided on my dissertation. I truly appreciate all the remarks mentioned in his review as they raised my awareness towards specific points that will surely enrich my dissertation. I will certainly endeavor to ensure that all the applicable remarks and comments are implemented in the final version or future publications. I will focus on the main and broad remarks in this review rebuttal due to time constraints and respond to some of the minor questions as the case might request.

First, I will respond to the reviewer’s remark about the double marking of vowel’s quality and duration that empower my claims that PA vowel inventory (and PA learners) seems to have a tense-lax rather than a pure length subdivision. This claim is motivated by the PA non-native listeners’ result in the perceptual assimilation model task in Chapter 3 of the dissertation. Historically, Arabic vowel inventory was never described to have a tenseness attribute attached to its different vowels in general. It was never introduced in the literature as such nor was it ever thought of as so. It is a moot point whether the spectral differences between the Arabic long and short vowels are large enough to qualify the contrast as tense-lax rather than just long-short. Only recently and for colloquial varieties of Arabic, other than Palestinian Arabic, such claims came up and were supported with scientific data. Some regional varieties of Arabic show larger spectral differences than others. For instance, Jordanian Arabic has spectral differences that are about as large as those found in English. In my dissertation, I merely charted this territory for listeners with Palestinian Arabic as L1. However, PA has not been studied in this respect yet. Therefore, any claims to a (budding?) tense-lax contrast are speculative. I never meant to say that Palestinian Arabic has replaced the short-long vowel contrast with a tense-lax contrast. The only thing my results permit me to say is that the short members of the vowels have more centralized positions in the quality space, and this property may have transferred to PA EFL.

Concerning the reviewer’s remark on the vowel length in AE being extremely unstable and, on the stimulus durations I used to test vowel production, I did not test vowels before fortis obstruents in my dissertation, and yet the non-native PA speakers of AE showed great overshooting of the produced AE vowels. Also, not only is there vowel shortening before voiceless obstruents, especially before voiceless stops, but there is also an extraordinary vowel lengthening before voiced obstruents. I have discussed matters in more detail in a follow-up paper (Van Heuven & Farran, 2022). Failure to apply the lengthening rule on the part of my PA EFL learners most likely explains the overall shortening of their vowels in English. There are quite a few pronunciation problems with AE vowels that I did not cover in my study. I am well aware that contextual influences (co-intrinsic vowel duration properties) will also be a source of difficulty. In light of the reviewer’s remarks on this issue and the related remark on Chapter 3, it can be hypothesized that the non-native learners would produce AE vowels with even shorter durations. This is a great suggestion for future research that I definitely would like to pursue.

Concerning the reviewer’s remarks on my presentation of the AE consonant system, the inventory I present in Table 2.6 was copied from Ladefoged (1999). In principle, the table lists consonant phonemes only. One exception, however, is the glottal stop but this is abundantly signaled as the odd man out. The double-starred note that accompanies the table explicitly mentions that glottal stop is NOT a phoneme (and thereby acknowledges that it should not be in the table). Possibly, I should have deleted the glottal stop from Ladefoged’s table. After all, it is an allophone of /p, t, k/, and it is not necessary for a nonnative speaker of AE to ever use the glottal stop.

This brings me to the criticism that it is unclear whether my dissertation deals with phonetic or phonological L1 interference. I am confident that both fields of linguistics are theoretically well-separated, but it is rather unmanageable to categorically state that certain realizations are purely phonetic or phonological. Therefore, I attempted to cover both areas in my dissertation as much as possible and provided the reader with sufficient information and background knowledge that may ease their understanding of the dissertation as one unit.

One can always mention additional peculiarities when reviewing a sound system of any language. My dissertation is not meant to be a comprehensive textbook on English phonetics and phonology; what I write about the consonants is merely part of a broad comparison between the consonant systems of Arabic and English. That being said, I would like to elaborate on a few minor remarks mentioned by the reviewer:

- Concerning the minor problems listed for Chapter 2 and specifically on the attributes of AE /h/, indeed, /h/ is always a classification problem. Basically, English /h/ is a voiceless vowel. Since it is a vowel, it has a well-marked formant structure and therefore also qualifies as a sonorant (like all vowels). The supra-laryngeal tract is excited by a noise source located at the glottis. Therefore, English /h/ is produced with a lot of friction but it is not a fricative since there is no obstruction in the supra-laryngeal tract. And yes, in retrospect, /h/ is better classified as a glide rather than a fricative. I shall be more critical towards the literature I used for the consonant inventory.
- Another minor issue brought up by the reviewer is the stability of obstruents' voicing. Indeed, voiced stops may be (partially) devoiced. However, this is an optional rather than a compulsory phenomenon. Whether the stops in words such as *obtain* and *Bob* are pronounced with or without vocal folds vibrating will not interfere with the speakers' intelligibility. What is important to maintain the voicing contrast is that voiceless stops are aspirated (in stressed syllables unless preceded by tautosyllabic /s/).
- Counter to what I write, the reviewer maintain that /ʒ/ may occur in the onset of English words, and cites the word *genre* as the example. My response is that *genre* is **not** an English word. It is a loan from French, as is also evidenced by the nasal vowel /ɑ̃/, which is not part of the English vowel inventory.
- Concerning the remarks on the /h, j/ not being part of any onset cluster and the examples that the reviewer mentioned, it should be highlighted that most phonological analyses consider the [j] not to be part of the onset but of the vocalic nucleus (turning long /u/ into a rising diphthong /ju/); the [j] sound only occurs immediately before /u/, and typically coalesces with the voiceless /h/ into voiceless palatal fricative [ç].
- The reviewer suggests transcribing the word *strengths* as [streŋgθ] or even [streŋkθ], i.e., with an intervening plosive between the *eng* and the fricative *th*. However, the *Cambridge Pronouncing Dictionary*, which I consulted when I needed the transcriptions, lists my /streŋθ/ - without the intervening plosive - as the only possible transcription for words like *strength* and *length*, both in American English and British English. <https://dictionary.cambridge.org/dictionary/english/strength>
- The reviewer comments on Page 26 of my dissertation that "...Primary and secondary stressed syllables are distinguished by the former for having higher pitch than the latter". While I appreciate the reviewer's remark on this matter, several amendments to that statement should be considered. The literature shows that it is not the height of the pitch that matters in English stress but the size and the synchronization of the pitch **change** (whether up or down). Primary stresses (which are indeed often realized with a prominence-lending pitch change) typically lose their pitch change when occurring in a non-focused unit, or when the pitch change is omitted to avoid stress clash. Secondary stresses, on the other hand, may optionally be given a pitch change, as in UNforGIVable, which then is not necessarily smaller than the pitch change on the primary stress. Moreover, secondary stresses are often given (much) larger pitch changes than the primary stress when the latter is omitted to avoid stress clash, as in predicative eighTEEN versus attributive EIGHteen YEARS. Presenting a concise comparison of AE stress vs. (Palestinian) Arabic stress patterns and rules was the main topic of my section on how stress is perceived.
- Another important remark mentioned under the section minor problems for Chapter 2 is related to one of my hypotheses on PA vowel perception, namely, that PA learners will confuse AE /ɪ/ with AE /ɛ/ but not with AE /e/. This prediction can be motivated as follows. The difference in length will be easier to perceive than the difference in vowel quality, especially when the learner's native language exploits duration for phonological vowel contrasts, and does not use fine-grained quality differences – such as Arabic. The duration ratio between /ɪ/ and /e/ is about 1:2 = 100% longer; the difference in vowel height between /ɪ/ and /ɛ/ is much smaller, ca. 35% (F1 frequency). **Accordingly, I predicted that, in the perception of PA EFL learners, /ɪ/ is closer to /ɛ/ than to /e/.**

- Concerning the reviewer's remark on page 28, half of the answer to the question proposed by the reviewer may be obtained from the results of the vowel identification experiment in Chapter 4. The 86 stimulus vowels were presented in a fixed /m_f/ context, i.e., before a voiceless consonant. The PA listeners almost exclusively used the duration cue and did not pay attention to the difference in vowel quality when choosing between the spectrally close long-short alternatives. But quite likely the PA EFL speakers will not respond like the native listeners when having to decide between spectrally adjacent long and short English vowels in variable contexts – just as they failed to lengthen the vowels before coda /d/ in the production task in Chapter 5.
- In response to the reviewer's question on Page 29, Table 2.10 of my dissertation on what I expect to happen to some English consonant allophones: e.g., the devoiced versions of /l, ɫ, j, w/ after aspirated voiceless stops, my expectation is that perception of these partly devoiced (i.e., aspirated) sonorants will be unaffected. Same for dark /l/. It will be readily perceived as an allophone of /l/ and the darkness will go unnoticed. However, in L2 speech production, the partial devoicing will be absent (although this will hardly interfere with the L2 speaker's intelligibility). The same goes for the pronunciation of dark /l/. Its absence will be noticed immediately by L1 listeners, but intelligibility will not suffer; there are also varieties of English that have no dark /l/, such as Irish English (or French-accented English). The Irish/French light /l/ is often made fun of but there are no indications that intelligibility suffers on account of it.

Concerning the minor questions and notes that the reviewer posited for Chapter 3, first on page 65 and the possibility of misinterpreting the informants' reactions. The suggested perspective is theoretically true, but the agreement and consistency I find in the participants' responses suggest that the overall results are sufficiently reliable to cover the entire group of EFL learners. Secondly, on page 70 concerning AE /ɑ/ *being in between* PA /a:/ and /u:/, *but closer to the former than to the latter*. My response is that AE /ɑ/ is definitely between PA /a:/ and /u:/, both in the degree of backness (F2) and in height (F1). Being in between /a:/ and /u:/ does not necessarily imply that AE /ɑ/ should be a point on a straight-line connecting PA /a:/ and /u:/. Rather it should be interpreted as: "In an area bounded by the coordinates of PA /a:/ and /u:/, AE /ɑ/ has a much shorter Euclidean distance to /a:/ than to /u:/."

As for the reviewer's general remarks on Chapter 4 and more specifically on the synthesized vowels, the point is that the same set of synthesized vowels was presented to native speakers of AE and the non-native PA learners. The results of this chapter, despite the mentioned confusions, still show good performance by the natives in identifying the vowels based on, not only the duration cue, but also the vowel quality. That is why we can see members of tense and lax vowels were chosen as responses for the different vowel durations presented in the task. Meanwhile, the non-native PA listeners fell short to achieve a competing performance as a response to the identical stimuli and highly depended on duration. Therefore, it is highly unlikely to submit to the idea that the confusions of the native listeners in this test were due to not considering any long vowel shortening. After all, the purpose of the task was to map out the different listeners' identification of the AE vowel inventory, not considering certain phonotactic constraints of the targeted language, which I am sure would result in very informative research in this domain and can be tackled in future research.

As for the minor questions and notes for Chapter 4, starting with page 84 concerning considering the shortening effect of the voiceless fricative on the preceding long vowels, and whether the (phonologically) long vowels were shortened accordingly in the stimuli. The answer is an affirmative yes, I did. The two durations chosen, 100 and 200 ms of vowel duration were tested with native listeners and found adequate for the particular context. These durations also correspond near-perfectly with the mean durations of AE short and long vowels before voiceless obstruents as reported by Lehiste & Peterson 1960.

As for the reviewer's remark on page 86 that the fixed carrier /CVI/ might tend to reduce the height of the tongue for the vowels because they are preceded by the dark /l/..., This is theoretically possible but unlikely. The listeners seem to have abstracted away from the effect of dark /l/, since the centroids of the 11 vowels in the reconstructed perceptual representation of the native listeners are exactly where they would be expected when compared with native speech production.

- Concerning the reviewer's remark on page 98 on a possible effect of the different dialects of the native speakers, the results of this task show that the dispersion of the vowel responses in L1 is a lot smaller than in the PA responses. More importantly, any vowel points near the phoneme boundaries will yield inconsistencies, even in the perception of native listeners. In the AE vowel system with 10 response categories (if we consider the low back vowels one category), the number of boundaries between spectrally adjacent categories is 24. The uncertainty generated by this large number of boundaries will account for much of the deviation from the modal response. Add to this the massive confusion between the merged low back vowels, and it can be realized that 60% agreement with the modal response is quite good.
- For page 102, the reviewer predicts that the real situation might be much worse when PA speakers have to produce and perceptually identify AE vowels in contexts that affect vowel length. This prediction will undoubtedly be confirmed, when tested in future research on PA learners. Therefore, the perceptual representation of the English sound system is probably even poorer than is shown by the current results. However, concerning the other remark for page 102 on the prominent confusions not involving the tense-lax contrast and the need for confusing examples to be used in the stimuli, I predict if the PA learners had picked up at least some of the effect of vowel quality (the way native listeners use it), the results would/might have been different.
- As for the reviewer remarks on page 113 H1, that probably would be the case, but this has to be established first in future experiments. In the mentioned hypothesis, I just draw conclusions based on the experimental data at hand.

Finally, as for the manuscript's terminological and clerical errors, I will surely take all the provided remarks into consideration. Again, I sincerely thank the reviewer for his time and efforts in reviewing my dissertation and for all the thought-provoking remarks and words of praise on my dissertation.

Budapest, August 23rd, 2022.



Bashar M. M. Farran

PhD candidate

Multilingualism Doctoral School, University of Pannonia