Arabic L2 Pronunciation in Australia
Preliminary observations on the production of emphatic /dˤ/ by students at The Australian University

Charbel El-Khaissi
The Australian National University
Canberra - Australia
charbel.el-khaissi@anu.edu.au

Abstract

While studies in the pronunciation of second language (L2) Arabic learners is an established field internationally, it is still maturing in Australia. This study is a preliminary investigation into the production of /dˤ/ and /d/ in the words /ˈbɑʕdæ/ ("after") and /ˈbɑʕdˤɑ/ ("some") by 15 L2 Arabic learners at The Australian National University, and their two native Arabic speaking instructors. A phonetic analysis was conducted in Praat to observe three pronunciation features relating to the articulation of /dˤ/ and /d/: intensity, voice onset time (VOT) and relative tongue placement and height via F1 and F2 formant values. The results demonstrate strong variation in the production of /dˤ/ and /d/, and although L2 learners do not make a distinction between pharyngealised /dˤ/ and its plain counterpart /d/, they adopt emphatic spreading on neighbouring vowels as a way to produce the necessary distinction. These preliminary findings in an Australian context are consistent with the literature on L2 speaker pronunciation of Arabic pharyngeal contrastive sets.

Keywords: Second-language learning, Arabic teaching, pharyngeals, Praat
1 Introduction

A well-known feature of Modern Standard Arabic’s (MSA) phonetic inventory is its four pharyngealised consonants /dˤ/, /sˤ/, /tˤ/ and /ðˤ/, which contrast against the non-pharyngealised sounds /d/, /s/, /t/ and /ð/ (Embarki, 2013, p. 31). The discrimination of Arabic pharyngealised consonants against their non-pharyngealised (‘plain’) counterparts is a difficult practice for Arabic language learners (Long and J. C., 2009; Shehata, 2018) and reflects a broader pronunciation challenge of distinguishing contrastive sound sets cross-linguistically (e.g., English /l/ and /v/ for Japanese speakers; Aoyama et al., 2004). As a preliminary study, this research project focuses on the production of one out of the four pairs of pharyngealised sounds, and specifically pharyngealised /dˤ/ and its non-pharyngealised counterpart /d/. The consonant /dˤ/ has been a great source of interest among scholars due to its confusion with other sounds in older Arabic texts, its gradual change or loss in modern Arabic dialects and because of the belief that this Arabic sound is unique among world languages to such an extent that Arabic is sometimes referred to as the language of daad (Versteegh, 1997, p. 89). The aim of this study is to conduct a contrastive phonetic analysis of Arabic L2 learners’ production of /dˤ/ and /d/, and compare it to native Arabic speakers. Specifically, the research questions this paper intends to answer are: how do L2 learners of Modern Standar Arabic at The Australian National University (ANU) differentiate between /dˤ/ and /d/ phonemes in their production, if at all? How does this compare to their native Arabic teachers? The general, exploratory nature of these research questions are fitting given the small data sample and stimuli, and aim to provide preliminary insights to support a more robust research investigation into L2 Arabic learner pronunciation in an Australian context, an area that remains mostly unexplored in the linguistic literature on L2 pronunciation.

2 Background

The /dˤ/ phoneme is a pharyngealised voiced, alveolar stop that is characterised by a primary and secondary method of articulation. The primary articulation typically involves the tip of the tongue being placed on the alveolar ridge, while the secondary method of articulation is characterised by the root of the tongue being retracted towards the pharynx, resulting in constriction of the pharyngeal wall. In principle, plain /d/ lacks this secondary method of articulation.

Image 1 (left) and Image 2 (right): Plain ص/s/ and emphatic ص/sˤ/ consonants, respectively (Shosted et al., 2018).

---

1 I presented an early version of this research project at The Australian National University’s Three Languages – Three Cultures: Narratives from the Middle East Conference (El-Khaissi, 2018). I thank all audience members who provided constructive feedback that helped develop the insights in this paper. Special acknowledgements are extended to Huda Al-Tammimi, Leila Kouatly, Alice Richardson, Hala Abu Alsisan, Turki Alharbi and the participants who volunteered in the study.
The above images pertain to the pharyngealised consonant /sˤ/ and its plain counterpart /s/ and are provided to visualise the nature of pharyngealised articulation. Notice in Image 2 the retraction of the root of the tongue to the back of the throat, resulting in constriction of the pharynx (indicated by the white circle); the back of the tongue also flattens to achieve this backed position. Meanwhile, for the plain consonant in Image 1, the root of the tongue is in its more relaxed, natural position with no constriction of the pharyngeal wall. This secondary method of articulation characterises the articulatory nature of all MSA pharyngeal consonants and is the source of production difficulty among second language Arabic learners, in addition to the degree of similarity between the contrastive sets which makes it difficult to make a productive distinction in L2 speech (Al Mahmoud, 2013).

Other articulation features worth noting with respect to pharyngealised sounds include intensity and voice onset time (VOT). As depicted in Image 3, pharyngealised consonants may display higher frequencies in intensity (or ‘emphaticness’) than their plain counterparts, while the VOT may be longer in plain consonants relative to their pharyngeal counterpart (Aldamen, 2013; Binasfour and Setter, 2018; Khattab, 2006), which reflects the general consensus in the literature notwithstanding dialectal variability.

A secondary phenomenon related to pharyngealised consonants that is the spreading of pharyngealised articulation to neighbouring vowels, otherwise known as progressive emphatic spreading or perseverative coarticulation (Shosted, Fu, and Hermes, 2018). This phenomenon involves the lowering and retraction of the tongue in the articulation of neighbouring vowels that precede or follow pharyngealised consonants. While some studies have revealed that the perception of pharyngealised consonants can be identified by the emphatic spread that takes place in adjacent vowels (e.g., Hayes-Harb and Durham, 2016; Jongman et al., 2011), there has been little to no attention given to whether emphatic spreading is used as a production strategy by L2 Arabic learners to discriminate between pharyngeal and non-pharyngeal pairs. This hypothesis will be treated later in the Findings section in relation to the first and second research question.

3 Literature Review

Shehata’s study of 45 native American English learners of Arabic (see Shehata, 2018) investigated the influence of proficiency levels—defined by years of Arabic instruction — on the perception and production of Arabic consonant clusters. His study targeted 18—30 years old students from a Midwest American university, in addition to three native Arabic speakers from the same university who listened
to participants’ recordings and rated their production in terms of correct or incorrect with high inter-rater reliability scores. Test stimuli included twenty monosyllabic Arabic nonwords (e.g., /daːk/ v. /dˤak/) and targeted ten Arabic contrastive sets, including the four aforementioned pharyngealised sets: /t-tˤ/, /d-dˤ/, /ð-ðˤ/, /s-sˤ/. The findings of Shehata’s investigation demonstrated that intermediate learners’ accuracy production was above chance for the /d-dˤ/ contrasts, while beginners performed at random chance for the same set. The production of /d-dˤ/, in addition to other sets, was marked by a higher variability among beginners, but this variation gradually converged among advanced speakers. Overall, the /d-dˤ/ set was ranked as having a high degree of difficulty in production relative to other consonants, as depicted in Figure 1 below.

![Figure 1: Percentage of difficulty of Arabic consonant phonemes in Shehata’s (2018) study (p. 66).](image)

A similar study by Binasfour and Setter (2018) recruited 38 female students of Arabic learners in Saudi Arabia with varying levels of proficiency, and aged between 20—26 years. Unlike Shehata’s experiment, the participants in their study were from both native English (UK, US) and non-native English (China, India) backgrounds. Production and perception tests included thirty Arabic words which filled the elicitation template ‘This word is …’ and targeted the Arabic emphatic consonants (C) /sˤ/, /dˤ/, /ðˤ/ and /tˤ/ across three pre-vocalic environments Cɑ Cʊ Cɨ via minimal pairs (e.g., /sˤar/ ‘became’ v. /sar/ ‘walk’). The results of this investigation revealed that both the perception and production difficulty of emphatic consonants was vowel dependent with Cɨ attracting the highest error rate, followed by Cʊ then Cɑ. These facts suggested that the vowel /ɑ/ may be more easily perceived and produced than the emphatics that precede the vowels /o/ and /i/. Furthermore, no significant differences were observed between language proficiency levels, although the authors noted this may be influenced by the unequal distribution of L1 speakers in each level in their study. Importantly, their findings indicate the strong effect of neighbouring vowels in the perception and production of Arabic emphatics—and in the case of emphatics preceding /ɑ/, it was evident in their study that L2 learners produced them more accurately (22.4% error rate) relative to /o/ (34.2%) and /ɪ/ (43.4%). This fact is supported by complementary studies in perception by Hayes-Harb and Durham (2016), which confirmed that native English speakers rely on adjacent vowels for cues to emphatic consonants. Further, they determine that transfer effects from the vowel space of native English speakers is responsible for the reduced accuracy of distinguishing emphatic consonants from their plain counterparts in the context of /o/ and /ɪ/. As schematised in their figure below, the English vowel space permits allophonic variation of /ɑ/ and /æ/—thus allowing for emphatic consonants and their plain counterparts to be functionally distinguished—while the same allophonic variation of /u/ and /i/ in plain and emphatic contexts does not exist in the same way. In effect, these results suggest that L2 Arabic learners are
able to more easily distinguish between emphatic and plain consonants if /ɑ/ occurs in the immediate vocalic environment.

4 Methodology

4.1 Participants

The participants in this research project were 15 native English speakers who were students of the Arabic language program at ANU in 2018, and their two native Arabic teachers. Nine out of the fifteen participants were female, while 6 were male. The age range of the speakers was 18—24 (mean=23.4). All participants were undertaking an intermediate or advanced Arabic language course at the time of the study meaning they had ample exposure and practice to pharyngeal phoneme production. Participants with a history of exposure to pharyngeal sound sets (e.g., heritage speakers of Arabic) or with competency in Semitic languages (e.g., Hebrew) were excluded. No participant had a history of language difficulty or impairment.

The two native speakers were Arabic language teachers at ANU. Their native dialects included Lebanese and Iraqi Arabic, however this is considered mostly inconsequential to the analysis due to the elicitation instructions provided to them (see §4.2 Procedure below).

4.2 Procedure

All participants were recorded with a studio-quality NT-USB Rode microphone in ANU’s soundproofed Speech Laboratory for optimal recording condition, and to ensure speakers’ privacy. A pop-filter was used on the microphone to minimise or eliminate any audio distortion caused by the impact of pharyngealised or plosive sounds. The microphone was mounted on an industry-standard tripod that allowed it to be at a comfortable height on a table-top at approximately 15cm from the speaker’s mouth. All speakers were initially instructed to speak into the microphone to test the best recording level and to ensure a standardised quality across the participants. The data was recorded in the software, Praat.

The MSA words overleaf were elicited:

The justification for the selection of the above two words was based on a variety of factors. Firstly, they are minimal pairs which differ only in the elements /dˤ/ and /d/, thus providing a stronger basis
for comparative analysis. Secondly, simple 2-syllable words are considered straightforward for intermediate and advanced students to pronounce and were verified by the Arabic program convenor and course teachers as having been explicitly taught or exposed to them during their language learning.

The duration of the recording sessions was approximately 15 minutes per participant. The native speakers were instructed to pronounce the words in MSA as they would in a language classroom setting, and not their respective Arabic dialects.

At the beginning of the recording session, each word was presented to the participant on a large placard. The participants had not been previously informed which word they would be pronouncing for the study and this helped to ensure the elicitation process reflected a spontaneous and natural pronunciation. Participants were given a moment to process the word before being prompted to read it out loud twice with a short pause in between each turn. If there was a disfluency or phonological error(s) in the speakers’ pronunciation, they were given the option to discard the recording and re-record.

4.3 Data Analysis

Recordings were analysed in Praat (Boersma and Weenink, 2009), an open-source acoustic analysis software for linguists, scientists and researchers. Three phonetic measurements were observed in this software, including formant values, intensity and VOT, and recorded in tabulated format. Formant extraction features in Praat approximate the place and manner of articulation by speakers, including the F1 value (approximating tongue height) and F2 value (approximating tongue location) 2

<table>
<thead>
<tr>
<th>speaker_id</th>
<th>vowel_id</th>
<th>context</th>
<th>f1</th>
<th>f2</th>
<th>f1_norm</th>
<th>f2_norm</th>
<th>f1_scaled</th>
<th>f2_scaled</th>
<th>iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>α</td>
<td>onset</td>
<td>574.6283</td>
<td>1439.0296</td>
<td>-0.929</td>
<td>-0.929</td>
<td>362.01</td>
<td>1196.437</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>α</td>
<td>onset</td>
<td>533.7963</td>
<td>1200.3058</td>
<td>-1.174</td>
<td>-1.877</td>
<td>332.759</td>
<td>850</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>α</td>
<td>end</td>
<td>801.6196</td>
<td>1486.2215</td>
<td>0.442</td>
<td>-0.742</td>
<td>524.911</td>
<td>1264.774</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>α</td>
<td>end</td>
<td>809.237</td>
<td>1831.3367</td>
<td>0.488</td>
<td>0.629</td>
<td>530.38</td>
<td>1765.792</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>ι</td>
<td>end</td>
<td>559.131</td>
<td>1821.5821</td>
<td>-1.022</td>
<td>0.59</td>
<td>350.832</td>
<td>1751.54</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>ι</td>
<td>onset</td>
<td>649.994</td>
<td>1892.7627</td>
<td>-0.473</td>
<td>0.873</td>
<td>416.112</td>
<td>1854.96</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>ι</td>
<td>end</td>
<td>949.6771</td>
<td>1831.9459</td>
<td>1.335</td>
<td>0.711</td>
<td>631.094</td>
<td>1795.758</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>ι</td>
<td>end</td>
<td>949.2013</td>
<td>1860.7137</td>
<td>1.332</td>
<td>0.745</td>
<td>630.737</td>
<td>1808.183</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>α</td>
<td>onset</td>
<td>533.8542</td>
<td>1308.5124</td>
<td>-0.735</td>
<td>-1.234</td>
<td>384.958</td>
<td>1084.978</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>α</td>
<td>onset</td>
<td>471.727</td>
<td>1446.9969</td>
<td>-1.048</td>
<td>-0.829</td>
<td>347.741</td>
<td>1232.981</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>α</td>
<td>end</td>
<td>974.211</td>
<td>1392.0865</td>
<td>1.482</td>
<td>-0.989</td>
<td>648.573</td>
<td>1174.511</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>α</td>
<td>end</td>
<td>794.88</td>
<td>1511.6476</td>
<td>0.579</td>
<td>-0.639</td>
<td>541.201</td>
<td>1302.415</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>ι</td>
<td>onset</td>
<td>584.9095</td>
<td>2137.0487</td>
<td>-0.478</td>
<td>1.193</td>
<td>415.517</td>
<td>1971.9</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>ι</td>
<td>onset</td>
<td>408.8532</td>
<td>2187.0495</td>
<td>-1.364</td>
<td>1.339</td>
<td>310.166</td>
<td>2025.255</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>ι</td>
<td>end</td>
<td>854.5334</td>
<td>1898.9416</td>
<td>0.879</td>
<td>0.495</td>
<td>576.873</td>
<td>1716.823</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>ι</td>
<td>end</td>
<td>815.7285</td>
<td>1956.9722</td>
<td>0.684</td>
<td>0.665</td>
<td>553.686</td>
<td>1778.948</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1: A sample of phonetic measurements of native and L2 speakers.

Formant extraction features in Praat approximate the place and manner of articulation by speakers, including the F1 value (approximating tongue height) and F2 value (approximating tongue location) 2

\[\text{VOT and intensity values are excluded here. The raw dataset is available open-source via Zenodo (El-Khaissi, 2021): http://doi.org/10.5281/zenodo.4451450}\]
The tongue trajectory of vowel articulation immediately following (non)-pharyngeal production was calculated via F1 and F2 formant value extraction at vowel onset and end point (see Image 4). Additionally, the VOT was measured in milliseconds for the consonant pressure build up time in addition to the intensity (measured in decibels) for the consonant release or burst (labelled ‘VOT’ and ‘burst’ in Image 4, respectively). The values for formants, intensity and VOT were averaged from the first and second recording iteration for each participant and native speakers then analysed in R-Studio (RStudio Team, 2020) using the ggplot2 package (Wickham, 2016). The vowels of all speakers were normalised with the Lobanov normalisation method to allow for inter-speaker comparison (Thomas and Kendall, 2007), and to eliminate any variation in formant values motivated by physiological differences between speakers (i.e., differences in mouth size or vocal tract). Normalised formant values were then scaled to hertz-like values for visual plotting. Normalisation and scaling procedures were conducted in R-Studio using the vowels package (Kendall and Thomas, 2018). Although sociolinguistic variables—like gender, age, and language level—featured in the participant questionnaire, they were excluded from statistical analysis due to data paucity, and to avoid overconfidence in the results.

5 Findings

In summary, the language learners’ pronunciation exhibited high variability. The following section summarises these findings in terms of intensity, VOT and vowel onset and midpoint trajectory following (non)-pharyngeal consonant.

5.1 VOT and Intensity among L2 learners

Language learners produced the non-pharyngeal /d/ phoneme with an average intensity of 62.7 dB, which was slightly lower relative to /dˤ/ 63.3 dB. This difference is more pronounced when the participants are categorised into two subsets according to whether they produce /dˤ/ with greater intensity relative to /d/ (indicated in red in Graph 1) or lower (in blue). Namely, eight speakers (53%) who produce /dˤ/ with less intensity relative to /d/ do so by -1.92 dB on average, while the seven speakers (47%) who produce /dˤ/ with a higher relative intensity to /d/ do so by +3.5 dB.

Regarding the VOT pronunciation feature (as in Graph 2), the majority of learners (13/15; 87%) produce /dˤ/ with a marginally longer VOT time of 9 ms relative to the production of /d/ which was 6 ms (-3 ms) on average.
5.2 Comparison of post-vocalic trajectory among L2 learners

On average, no significant distinction was observed between the articulation onset of normalised vowels following pharyngeal /dˤ/ (F1: 461 Hz; F2: 1668 Hz) and non-pharyngeal /d/ (F1: 416 Hz; F2: 1878 Hz), as depicted visually in Graph 4. While the /dˤ/ phoneme was slightly more retracted relative to /d/ (F2: -210 Hz) and lowered (F1: -45 Hz), such differences are considered too minuscule to be functionally distinct when benchmarked against the two native speakers in this study (see Graph 5 for a comparison of native and L2 speakers).

More significant differences were observed in the endpoint trajectory of the same phonemes. While the tongue height (F1) was almost identical in the endpoint across the phoneme set (/dˤ/: 503 Hz; /d/: 509 Hz) the tongue was further retracted (F2) on average in the production of pharyngeal /dˤ/ (1135 Hz) relative to non-pharyngeal /d/ (1463 Hz). Regarding L2 speaker variation, learners appear to vary more in terms of the final tongue height placement in the production of /d/, while more variation was observed in the final tongue backness in the production of /dˤ/.

5.3 Comparison between L2 learners and native speakers

In summary, no significant differences were observed in the acoustic features of intensity and VOT among native speakers and L2 learners. Regarding intensity, native speakers produced the non-pharyngeal /d/ phoneme with an average intensity of 51.1 dB, which was slightly lower relative to /dˤ/ (53.1 dB). The +2.0 dB discriminatory difference in the intensity of /dˤ/ production relative to /d/ in native speakers is similar to the subset of L2 learners who produced /dˤ/ with a higher relative intensity to /d/ (+3.5 dB).

Regarding the VOT pronunciation feature, native speakers on average produced no difference in the VOT time between /dˤ/ (8 ms) and /d/ (8 ms), which contrasts against L2 speakers where a marginally longer VOT time (+3 ms) was observed in the production of /dˤ/ relative to /d/.

Significant differences were observed in the tongue trajectories of vowels immediately following /dˤ/ and /d/ production between native speakers and learners (Graph 5). Beginning with /d/ (indicated in red), there is a clear lowering and retraction of the tongue among L2 learners and while this mimics the
**Graph 3**: Approximate L2 tongue location at vowel onset following pharyngeal (PH) /ḍ/ and non-pharyngeal (NPH) /ḍ/ groups, including average midpoint (large circle) and individual points (small circle).

**Graph 4**: Approximate L2 tongue location at vowel end point following pharyngeal (PH) /ḍ/ and non-pharyngeal (NPH) /ḍ/ groups, including average midpoint (large circle) and individual points (small circle).
general trajectory or direction of tongue movement among native speakers, there is clearly a greater retraction of the tongue approximating /ɑ/ among L2 learners, as opposed to a more fronted /æ/ as in native speakers\(^3\). In terms of the vowel following /dˤ/ (indicated in blue), L2 learners on average demonstrate a significant retraction of the tongue in the articulation of the following vowel: /dˤɑ/. Interestingly, the vowel-ending is characterised by the retraction of the root of the tongue further back in the throat than /dˤ/ onset in native speakers. Conversely, the trajectory of post-vocalic /dˤɑ/ among native speakers proceeds in the opposite direction to L2 learners, that is, a clear fronting from the back of the throat towards /æ/.

6 Discussion

The results of this study, albeit preliminary in nature, provide interesting insights into pronunciation features of Arabic L2 learners in an Australian context that can be used as a basis for further investigation, and also stimulate discussion on L2 Arabic pronunciation strategies more generally.

Beginning with the first acoustic parameter, intensity, nearly half of the participants produce /dˤ/ with a higher relative intensity to /d/ by +3.5 dB on average. Although there is no unanimous agreement regarding the minimum change in intensity that produces a perceptual difference in Arabic pharyngeals, the native speakers’ average intensity increase of +2.0 dB for the pharyngealised consonant may be used as a benchmark against this subset of learners to suggest that intensity may be used by some L2 learners as an acoustic cue for pharyngeal set discrimination; similar findings were observed in Binasfour (2018, p. 376) for the set /s/-/sˤ/. However, when we consider this subset alongside the

---

\(^3\)See Appendix for individual trajectories.
remaining L2 learners (53%) whose intensity of /dˤ/ is lower relative to /d/ (-1.92 dB), then the preliminary picture is one of an almost equal chance, suggesting strong variability in the use of intensity as a discrimination mechanism for /dˤ/-/d/ among the L2 speakers.

Regarding the VOT pronunciation feature, the average difference of 3 ms between /d/ relative to /dˤ/ is considered too insubstantial to suggest VOT to be a reliable acoustic cue for pharyngealization discrimination, unlike other studies where it was found to be noticeably longer in plain consonants (Binasfour, 2018; Khattab, 2006). Still, the single outlier with a VOT difference of 13 ms (see Graph 2) reflects that some L2 Arabic learners produce /dˤ/ and /d/ with a longer VOT similar to existing studies (e.g., 16 ms average difference in Aldamen, 2013) and thus reinforces the overall variability among student pronunciation in classrooms.

The near-identical formant values in the vowel production immediately following /dˤ/ and /d/ suggests that, on average, L2 Arabic learners do not meaningfully discriminate between the consonant set; or rather, any meaningful attempt to distinguish between the production of /dˤ/ and /d/ was unsuccessful, and these preliminary observations are consistent with previous studies that similarly observed a low discrimination accuracy between the same contrastive consonant sounds, including Shehata (2018).

The supposed misarticulation of /dˤ/ and /d/ consonants does not necessarily mean L2 learners do not discriminate between words containing /dˤ/ and /d/. The noticeable ‘backness’ of the tongue in the vowel endpoint for /dˤɑ/ relative to the more fronted /dæ/ appears to be clear evidence in support of emphatic spreading. In particular, the vowel following /dˤ/ among L2 learners reflects an articulation with the tongue lowered and retracted, which is characteristic of pharyngealised consonant qualities. In fact, the L2 learners’ articulation of this vowel exhibits stronger pharyngealisation effects compared to native speakers. This suggests L2 learners are capable of achieving a target tongue position in the same space as where native Arabic speakers produce pharyngeals, contrary to misconceptions that this is a physiological barrier to non-native Arabic speakers. Further, and more interestingly, such a result raises the question of whether L2 learners use emphatic spreading in neighbouring vowels as a compensatory acoustic strategy to counteract the lack of production discrimination between /dˤ/ and /d/ consonants. Indeed, perception studies among native Arabic speakers have demonstrated that pharyngealised consonants in words are not necessarily identified by the articulation of the pharyngealised consonant itself, but by the pharyngealised qualities which are taken up by adjacent vowels (e.g. Jongman et al., 2011 in Jordanian Arabic). From a pedagogical point of view, this raises an issue in terms of whether native Arabic teachers are positively reinforcing students’ discrimination of pharyngeal and non-pharyngealised consonants via the presence of emphatic spreading and not the articulation of the consonant itself. In fact, native English speakers have been suggested to rely more on following vowels than on the consonants when discriminating Arabic emphatic and plain consonants, and this is especially true for /ɑ/-/æ/ vowel contrasts (e.g. Hayes-Harb and Durham, 2016). Importantly, the fact that L2 learners adopt emphatic spreading in their articulation of the vowel following /dˤ/ suggests they are capable of articulating pharyngealization, and thus any perceived misarticulation is not necessarily an inability for L2 learners to retract the tongue in order to achieve a primary constriction in the pharynx, but rather the execution of this method upon consonants specifically.

7 Conclusion

This study explored the acoustic qualities of Arabic consonant /dˤ/ contrasted against its plain counterpart /d/ among L2 Arabic learners at The Australian National University. Overall, the findings affirmed two key characteristics of L2 pronunciation among English speakers in an Australian context that is consistent with the literature: (1) the production discrimination of pharyngealised consonant sets, and in this case /dˤ/ and /d/, is highly variable among students, and; (2) F2 formant values, relating to
tongue backness or retraction, remains the most reliable acoustic cue for differentiating between such contrastive sets, and particularly on neighbouring vowels as opposed to the consonants themselves.

The study adds to the growing body of literature examining L2 learners’ acquisition of Arabic emphatics, and has two pedagogical implications. First, the findings provide evidence of clear emphatic spreading on vowels following /dˤ/, which reinforces recommendations to direct learners’ attention to neighbouring vowels when producing Arabic pharyngeals; existing Arabic language textbooks, including the popular Alif Baa: Introduction to Arabic, provide some indication of this already (Brustad et al., 1995 as cited in Hayes-Harb and Durham (2016, p. 562), but students would no doubt benefit from instruction that drew attention explicitly to surrounding vowels, including vowel type effects. Second, there are clearly a number of pedagogical benefits to being able to visualise tongue height and placement during the articulation of speech sounds by language learners. However, Praat’s non-user-friendly interface and complex functionalities are often cited as reasons against its usage in a language learning context by language learners and teachers (e.g. Brett, 2004), and while it supports formant extraction processes, it offers little in the way of visualisation. For this reason, more computational effort is required to build classroom-friendly software for the automation of formant extraction procedures and their visualisation (e.g., see prototype by Chao et al., 2020).

Since this study is among the first in Australia to investigate the production of emphatics by L2 Arabic learners, there is no doubt a need for future investigation of this kind, including ones with larger participant sample sizes and more comprehensive elicitation tasks that cover the full range of Arabic pharyngeal contrastive sets and in various vocalic environments (e.g., /i/, /u/). In addition, the present study excluded sociolinguistic variables to understand whether aspects such as proficiency level (beginner, intermediate, advanced), exposure to Arabic outside classroom or time abroad in Arabic-speaking countries condition the discrimination of contrastive sets. For instance, Shehata’s study (see Shehata, 2018) indicated that learner production performance improved from 57% for beginners to 63% for intermediate to 93% for advanced learners—but study abroad experience and direct exposure to the Arabic language did not appear to play a significant role in the production (and perception) accuracy given that Shehata’s learners had never studied Arabic abroad or spent extended periods of time in an Arabic-speaking country (p. 60). Furthermore, future studies of this kind would benefit from a larger number of native speakers, including across multiple native Arabic dialects, in order to create a stronger comparative basis with L2 learners, but also to investigate the degree by which native dialect phonology plays a role in MSA pronunciation. While outside of scope for the present study, the same investigation of native dialect phonology would apply, at least in principle, to Australian participants recruited in future research, and whether their native English dialect has any significant effect on the quality of their Arabic pronunciation relative to native English speakers of other dialects, such as British or American English. Hayes-Harb and Durham (2016), among other authors, have already proposed the probable existence of transfer effects in the mapping of Arabic vowel allophones to the English vowel space (recall Figure 1), but further investigation is required to understand if this is also dialectally motivated.
Appendix

Individual post-consonant vocalic trajectory of pharyngeal /dˤ/ (‘PH’) in Graph 10 and its plain counterpart (‘NPH’) in Graph 9 using non-normalised values. The five speakers with a positive NPH trajectory (indicated in blue) are the same speakers with positive trajectories for PH (with the exception of one additional person). While outside the scope of the present study, it is included here as a note for further investigation since it reflects the possibility of articulation errors being generalised between contrastive sound sets.

Graph 9: Tongue trajectory during articulation of vowel following non-pharyngeal + (NPH) among learners (L2).

Graph 10: Tongue trajectory during articulation of vowel following pharyngeal + (PH) among learners (L2).
References


