Acoustic Properties of Foreigner Directed Speech

Azza Al Kendi[1], Ghada Khattab[2]

1. School of English Literature, Language and Linguistics, Newcastle University, UK
2. Speech and Language Sciences, Newcastle University, UK
A.N.S.Al-kendi2@newcastle.ac.uk, ghadah.khattab@newcastle.ac.uk

ABSTRACT

The acoustic properties of vowels in clear speech have been extensively investigated, but a less explored area is that of foreigner directed speech (FDS). This study examined the production of Arabic /iː/, /uː/, /aː/ in the speech of 22 native Omani speakers, either addressing their foreign domestic helpers (FDH) or a native speaker (NS). Words containing the target vowels were elicited using an interactive spot the difference task. Vowel space, F1 and F2 measures, F0, intensity and duration were compared in productions to FDH and NS interlocutors. Speech to FDHs yielded higher F1, greater vowel space expansion and higher pitch and intensity than speech directed to the NS, but the differences were modest and there was no effect on F2. Moreover, vowels in speech to FDHs were surprisingly shorter in duration than those in speech to the NS. External factors relating to the peculiarity of the FDH context are considered.

Keywords: foreigner directed speech, acoustic properties

1. INTRODUCTION

Talkers tend to speak more clearly and to hyperarticulate their speech sounds under certain conditions such as a noisy environment or when interacting with listeners who are deemed to require more intelligible speech; these include hearing-impaired listeners, infants/children and non-native listeners. The acoustic modification of vowels has been widely investigated in clear speech and infant directed speech (IDS) research.

Clear speech is typically characterized by an expanded vowel space compared to conversational speech [5, 8, 19], though this is not immediately obvious in all the acoustic patterns investigated. For instance, F1 is typically found to be higher in clear speech regardless of vowel category, indicating that speakers produce more open vowels and increase their vocal effort [8]. On the other hand, F2 patterns do indicate expansion. Clear front vowels have typically been found to have higher F2, while clear back vowels have lower F2, maximising the area between front and back vowels [8]. With respect to duration, vowels in clear speech are generally longer in duration than those in conversational speech [5, 9, 19]. These hyper-articulated properties of clear speech are not unexpected since clear speech is elicited in contrived laboratory settings using linguistically controlled stimuli.

IDS has been similarly shown to exhibit an expanded vowel space that is believed to have a didactic role due to infants' linguistic needs [13, 16]. However, a study by [6] did not find stretching of the vowel triangle along the F1/F2 dimension in IDS, but rather a shift of the vowel triangle along the F1 dimension, suggesting more open vowels. IDS is also characterized by a slower speech rate and higher pitch [13, 20]. Heightened pitch contours and varied pitch ranges in IDS are believed to arouse a child's attention and enhance positive affect [10]. FDS research has hypothetically assumed that FDS should be similar to IDS or clear speech given that foreigners have linguistic needs in the target language [22]. Indeed, a small body of research found FDS to have a larger vowel space and longer vowel duration, but not higher mean f0 compared to adult directed speech (ADS) [20, 22]. This was explained as demonstrating that FDS shares the same properties as IDS when it comes to the didactic role, but lacks increased properties of affect.

Differences in the methodologies used and the heterogeneity of FDS contexts point to the need for further research in this area. For instance, apart from the issues relating to F1/F2 expansion highlighted above, it is unclear whether prosodic position has been controlled for. Further, some studies on FDS used contrived data collection tools such as hypothetical listeners to elicit speech samples. Lastly, vowel intensity and loudness have not been examined directly in the abovementioned research.

The current study extends existing research on FDS and examines speech to an understudied but significant population of 'foreigners', that of African and Asian domestic helpers in the Gulf. The study focuses on the acoustic characteristics of Arabic vowels in FDS, expanding this area of research beyond work on English; it also provides a first report on vowel intensity in FDS. As these foreign helpers originate from different countries and have different first languages, the little research that there is has so far focussed on their own pidgin form of Arabic rather than speech addressed to them [2].

Based on the above-mentioned studies, we predict that vowels in FDS would have an expanded vowel space and longer duration. We also hypothesize that FDS and ADS would not be different in f0 but that FDS would have higher intensity.

2. METHOD

2.1. Participants

To elicit FDS, 22 female native Omani speakers (NSs) (mean age: 34) and their female foreign domestic helpers (FDHs) participated in the study. The FDHs were diverse in terms of their country of origin, length of residence (LoR) and Arabic proficiency level. They came from countries in Asia and Africa (e.g. India, Philippines, Bangladesh, Sri Lanka, Indonesia, Tanzania, Ethiopia, Nigeria, and Uganda). They had varying Arabic experiences based on their LoR in the Arab world (0.7- 21 years (mean: 6.23). They had been exposed to Arabic in a naturalistic setting and had not received formal instruction in the language. They all spoke with a noticeable foreign accent but their proficiency level in Arabic varied based on their LoR in the Arab world. Thus, those who had been in the Arab world for quite a long time could communicate better than those who had spent a few months or years. They had been working with their current employers for no less than two months (0.16 - 4 years). As a comparison group and to elicit ADS, a female native Omani adult was recruited (the first author). She was from the same home town of the NSs and spoke the same dialect. The participants reported no history of hearing problems.

2.2. Material

Nine content Arabic words that contain one of three Arabic long vowels (/iː/, /aː/, /uː/) (/fiːl/ ‘elephant’, /tiːn/ ‘fig’, /ħaliːb/ ‘milk’, /baːb/ ‘door’, /taːg/ ‘crown’, /kɪtaːb/ ‘book’, /fuːl/ ‘chickpeas’, /tuːt/ ‘berries’, /χaruːf/ ‘sheep’) were used. All vowels appeared in a stressed syllable.

2.3. Procedure

In order to obtain comparable samples of the target vowels across speech registers, participants were engaged in a spot the difference task [1]. The task consisted of six picture pairs with three different scenes, two pairs per scene, one to elicit FDS and the other to elicit ADS. The scenes had objects that represented the target words and some distracters. In two consecutive sessions, each participant was instructed to sit opposite her interlocutor and to try to spot twelve differences between each picture pair without seeing each other’s pictures. The NS had all the missing items on her version of the pictures and was encouraged to take the lead and negotiate the differences with her interlocutor. The interaction was recorded using Edirol digital recorder R-09HR by Roland with a sampling rate of 44.100 Hz and 16-bit quantisation. This was connected to a Sennheiser radio microphone which the participant wore.

2.4. Acoustic Analyses

A total of 680 words were acoustically analysed in both conditions (FDS=429 & ADS=251). Impressionistic examination revealed that all words appeared in a stressed position in the utterance. Acoustic measurements were carried out in Praat using a script. Formant frequencies were manually checked for any errors that might result from automatic extraction.

The vowel space area represented the triangular area encompassing the three vowels constructed from the averaged F1 and F2 values of both FDS and ADS conditions on the x-y plane [14] using the PhonR package [17]. F1 and F2 values were obtained from vowel midpoint. They were converted to the psychoacoustic Bark scale [21]:
$Z = \{26.81/(1+1960/f\}-0.53$, where Z is the critical-band value of a formant in Bark and f is a formant’s frequency in Hertz. Duration was calculated from the waveform and spectrogram. Vowel F0 and intensity were obtained from vowel midpoint.

2.5 Statistical Analysis

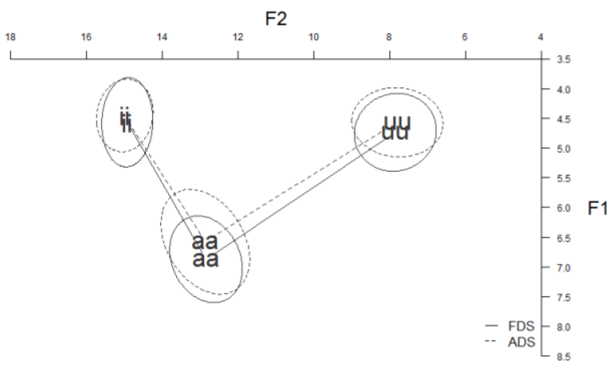
Statistical analyses were carried out in R. The polygon area that connects vowel formant means was used to obtain vowel space means using phonR. To obtain p-values for vowel space area, a simple t-test was used. Linear mixed effect models were used to test the effect of speech register on F1, F2, duration, F0 and intensity using lme4 package [4]. Speech register was used as fixed effect. As random effects, we had intercepts for speakers and lexical items. To test whether any modification in F1, F2, duration, F0 and intensity is the by-product of prosodic position in which the vowels appear, each vowel was assigned either 0 or 1 depending on whether the vowel appeared in a word at a phrase boundary or not [18]. Vowels that appear at a phrase boundary might be longer in duration, which could contribute to F1/F2 expansion. A full linear mixed effect model including speech register and prosodic boundary as fixed effects was compared against a reduced model without speech register for all three vowels (10 models). The random effect structure was similar to the one above. P-values were obtained by likelihood ratio tests [3].

3. RESULTS

3.1. Vowel Space Area

A simple t-test revealed a significant main effect of speech register (t= 33.7, df=1, p= 0.01). The overall mean vowel space area of FDS (7.9 Barks) was larger than that of ADS (7.4 Barks) (Figure 1). However, the expansion of the vowel space area in FDS was driven by a change in the F1 rather than the F2 dimension.

Figure 1: Vowel space of Arabic vowels in FDS and ADS



3.2. F1 & F2 at Steady State

For F1, LMER revealed a significant effect of FDS on the increase in F1 for all vowels ($p<0.01$). F1 of FDS was higher by 0.11 Barks ± 0.07 (standard errors) for /i:/, 0.2 Barks ± 0.09 (standard errors) for /a:/, 0.15 Barks ±0.04 (standard errors) for /u:/ compared to ADS. Likelihood ratio tests revealed that speech register contributed significantly to the increase in F1 for all vowels and beyond the effect of prosodic boundary (χ^2 (1)= 6.65, $p<0.01$, ($\chi^2(1)=20.58$, $p<0.01$, $\chi^2(1)=11.32$, $p<0.01$ for / i:/, / a:/, /u:/ respectively).

For F2, LMER revealed no significant effect of speech register on /i:/ or on /u:/. F2 for /i:/ in FDS was lower by 0.08 Barks ± 0.09 (standard errors), $p>0.05$ compared to that in ADS. F2 for /u:/ in FDS was higher by 0.09 Barks ± 0.09 (standard errors), $p>0.05$.

3.3. Vowel Duration

LMER revealed a significant difference between FDS and ADS in regard to vowel duration. Vowels in FDS were found to be shorter by 9.41 ms ±2.82 (standard errors), $p<0.01$ than vowels in ADS. Further analyses were carried out to find out if longer vowel durations in ADS are due to prosodic boundary. A likelihood ratio test revealed that the full model with the effect in question (speech register) was significant ($\chi^2(1)=13.33$, $p<0.01$), indicating that the effect of speech register is more prevalent than that of prosodic boundary. Another factor that could contribute to a reduction in word duration is word repetition, especially in spontaneous speech [12]. In our case, there were more word repetitions in FDS than in ADS, which could have contributed to this reduction in vowel duration in FDS. Thus, we excluded all word repetitions from both registers. LMER revealed that the difference in vowel duration between both registers was now insignificant ($p>0.05$). Vowels in ADS were still slightly longer but only by 3.1 ms ± 3.43 (standard errors).

3.4. Vowel Intensity and F0

LMER revealed a significant effect of speech register on vowel intensity ($p<0.01$). Vowels in FDS were higher in intensity by 1.72 dB ±0.28 (standard errors) than those in ADS. A likelihood ratio test showed that speech register contributes significantly to this increase in vowel intensity and beyond the contribution of prosodic boundary ($\chi^2(1)=33.55$, $p<0.01$).

For vowel F0, there was a significant effect of speech register ($p<0.01$). Vowels in FDS were higher in F0 by 16.42 Hz ± 2.22 (standard errors) than those in ADS. A likelihood ratio test revealed a more significant weight for the full model with the effect in question (χ^2 (1)=52.44, $p<0.01$).

6. DISCUSSION

The results of this study shed light on the acoustic properties of vowels in speech directed to a special group of foreigners learning the target language in a naturalistic setting. F1 was significantly modified in FDS compared to ADS. The increase observed in F1 for the three vowels proved to be mainly the effect of speech register rather than the prosodic boundary in which the vowels appear. F2 of the front vowel was slightly lower in FDS than that of ADS and F2 of the back vowel was slightly higher in FDS than that of ADS though differences were hardly noticeable and insignificant. The increase of F1 in FDS is consistent with previous findings on IDS and clear speech [6, 8, 9]. It indicates that speakers are producing more open vowels as they increase their vocal effort. Vowel space was found to be expanded in FDS compared to ADS. This pattern is in line with previous research [8, 13, 22]. However, findings of vowel space in the current study reflect changes in F1 but not F2, which is against the hyperarticulation hypothesis.

To address the lack of FDS effect on the front vowel /i:/ or the back vowel /u:/ with regards to F2 and more generally the lack of support for the hyperarticulation hypothesis, we may refer to the special FDH context. The hyperarticulation of speech sounds observed in clear speech are not surprising given the methodological conditions set for these studies. Hence, a more peripheral position of these vowels may be more evident if speech was constrained by the linguistic context or speech was given in other conditions similar to these used to elicit clear speech. In addition, hyperarticulation in IDS has been challenged [6, 7], and the linguistic needs of infants are different from those of adults learning a foreign language. Furthermore, NSs in this study may not have seen a need to hyperarticulate their speech to FDHs due to the fact that NSs in this study were more familiar with their FDHs than they were with the NS interlocutor. To this end, this finding highlights the importance of examining F1 and F2 of each vowel separately before making conclusions about any expansion in vowel space as well as considering the context of study. From this, we cannot conclude that modifications in vowels in speech directed to FDHs entail hyperarticulation or serve a linguistic benefit for the language learners.

The results also showed that vowels in FDS were significantly shorter than those in ADS. This finding was against our prediction and was found to be inconsistent with previous findings [5, 7, 8, 13, 20]. This raises the issue of why vowels in FDS are shortened. We considered word repetitions in FDS, since the task lent itself to spontaneous interactions in which speakers had the chance to repeat words for clarity of context or further negotiation. This was indeed shown by the larger number of words addressed to FDHs than to the NS. Findings revealed that the difference in vowel duration between ADS and FDS was insignificant as it dropped from 9.48 ms to 3.1 ms when repeated words were removed from analysis. Hence, repeated words is one way to explain the significant shortening of vowels in FDS. This once again puts into question the didactic or clear speech aspects of FDS to this population of learners. Another likely explanation for this finding is the degree of familiarity with the subjects. Further research is needed to assess the relationship between familiarity with the subjects and acoustic modifications in speech.

Mean fundamental frequency (F0) or pitch was also found higher in FDS compared to ADS. This was mainly due to an effect of speech register and not a by-product of prosodic position. This finding was not in line with FDS studies, which did not find a difference between FDS and ADS with regards to mean F0 or pitch range [20, 22]. This finding is, however, consistent with IDS findings though differences of F0 between IDS and ADS are more robust [22]. To explain this deviation from previous FDS studies, we may consider the NS-FDH context and compare it to that of the mother-child context. If mothers use attention-eliciting cues by heightening their pitch when talking to their children, native speakers might similarly arouse their FDH's attention to the target words by increasing word F0. This might be especially valid given the similar language learning setting of infants and FDHs. Research has also shown that increase in pitch is sometimes caused when mothers repeat target words in response to a child's failure to attend properly [11]. Given that high F0 is a property of auditory signals intended to be alerting, elevating pitch on consecutive repetitions could be an effective strategy to call the child's attention to the target word [11]. Similarly, NSs in this study could have used an equivalent strategy given that NSs repeated themselves when FDHs failed to comprehend the target word or the speech context. Future research should investigate this further by examining similar FDH contexts.

Vowel intensity significantly increased in FDS compared to ADS. This was primarily due to an effect of speech register and not a consequence of prosodic boundary in which the vowels appeared. Comparisons of this finding with other work on speech register are hard to make due to a lack of research on the effect of speech register on vowel intensity specifically. An increase in a sound's intensity is the result of an increase in respiratory effort [15]. Sounds produced with higher intensity are perceived as louder by listeners. Although not as powerful as F0, intensity is one prosodic cue to linguistic emphasis [11]. The FDH context might be one factor triggering this increase in intensity. Exaggerated intensity rates or speaking louder could therefore be a strategy used by NSs to emphasize target words. Further research is needed to assess this property.

Taken together, these findings show how changes in speech directed to special listeners are unique to the listeners' context. The results reveal modest expansion in the way the vowel space was larger in FDS and F1 was higher. However, the lack of F2 difference between FDS and ADS and the significant shortening of vowels in FDS put the hyperarticulation hypothesis into question. The NSs familiarity with their FDHs might have triggered the lack of FDS. The results also reveal that F0 and intensity were higher, which makes it likely that these prosodic cues are being used to emphasise target words and stimulate auditory attention to FDHs.

7. REFERENCES

[1] Baker, R. and Hazan, V. 2011. DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behaviour Research Methods*, *43*(3), 761-770.

[2] Bakir, M.J., 2010. Notes on the verbal system of Gulf Pidgin Arabic. *Journal of Pidgin and Creole Languages*, *25*(2), 201-228.

[3] Barr, D., Levy, R., Scheepers, C. and Tily, H. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

[4] Bates, D, Maechler, M., Bolker, B. and Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.

[5] Bradlow, A, Torretta, G. and Pisoni, D. 1996. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, *20*(3-4), 255-272.

[6] Dodane, C. and Al-Tamimi, J. 2007. An acoustic comparison of vowel systems in adult-directed-speech and child-directed speech: Evidence from French, English & Japanese. In *16th International Congress of Phonetics Sciences*, 6-10.

[7] Englund, K.T. and Behne, D.M., 2005. Infant directed speech in natural interaction—Norwegian vowel quantity and quality. *Journal of Psycholinguistic Research*, *34*(3), 259-280.

[8] Ferguson, S., Kewley-Port, D. 2002. Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *J. Acoustic Society of America,* 112, 259-271.

[9] Ferguson, S. H., Kewley-Port, D. 2007. Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *J. Speech Language Hearing Research,*. 50, 1241-1255.

[10] Fernald, A. and Kuhl, P. 1987. Acoustic determinants of infant preference for motherese speech. *Infant Behavior and development*, *10*(3), 279-293.

[11] Fernald, A. and Mazzie, C., 1991. Prosody and focus in speech to infants and adults. *Developmental Psychology*, *27*(2), 209.

[12] Fowler, C.A. 1988. Differential shortening of repeated content words produced in various communicative contexts. *Language and Speech*, *31*(4), 307-319.

[13] Kuhl, P., Andruski, J., Chistovich, I., Chistovich, L., Kozhevnikova, E., Ryskina, V., Stolyarova, E., Sundberg, U. and Lacerda, F. 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*(5326), 684-686.

[14] Ladefoged, P. 1996. *Elements of Acoustic Phonetics*. Chicago: University of Chicago Press.

[15] Lehiste, I. and Lass, N.J. 1976. Suprasegmental features of speech. *Contemporary Issues in Experimental Phonetics*, *225*- 239.

[16] Liu, H., Kuhl, P. and Tsao, F., 2003. An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6*(3), F1-F10.

[17] McCloy, D. 2016. PhonR: tools for phoneticians and phonologists. R package version 1.0-7.

[18] Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H. and Mazuka, R. 2017. Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition*, *166*, 84-93.

[19] Picheny, M., Durlach, N., Braida, L. 1986. Speaking clearly for the hard of hearing. II: Acoustic characteristics of clear and conversational speech. *J. Speech Language. Hearing Research*, 29, 434–446.

[20] Scarborough, R., Dmitrieva, O., Hall-Lew, L., Zhao, Y. and Brenier, J. 2007. An acoustic study of real and imagined foreigner-directed speech. *Journal of the Acoustical Society of America*, *121*(5), 3044.

[21] Traunmüller, H. 1990. Analytical expressions for the tonotopic sensory scale. *J. Acoustic Society of America*, 88, 97–100.

[22] Uther, M., Knoll, M.A. and Burnham, D. 2007. Do you speak E-NG-LI-SH? A comparison of foreigner-and infant-directed speech. *Speech Communication*, *49*(1), 2-7.