Can you hear me? Acoustic modifications in speech directed to foreigners and hearing-impaired people.

Monja Angelika Knoll, Melissa Johnstone, Charlene Blakely

Applied Psychology Research Group, The University of The West of Scotland, Scotland, UK

Monja.knoll@uws.co.uk, B00216756@studentmail.uws.ac.uk, Charlene_7799@hotmail.com

10.21437/Interspeech.2015-618

Abstract

People modify their voice to accommodate the communicative needs of their audience. Previously, speech recipient groups with communicative needs such as foreigners and infants have been investigated, however, there is a lack of research into speech directed to hearing-impaired people. The aim of our study was to investigate prosodic modifications in speech directed to hearing-impaired people and foreigners, and the benefit of such modifications to the listener. Ten speakers were recorded in natural interactions with a hearing-impaired person (HIDS), a foreigner (FDS) and a native speaker (ADS). The resulting speech was acoustically analysed for duration, vowel hyperarticulation, intensity, mean fundamental frequency (F_0) and F_0 range. There was no significant difference in duration, F_0 range or intensity between the conditions. HIDS and FDS displayed significantly higher mean F_0 and hyperarticulation than ADS. The speech was also rated by twenty speakers for positive vocal affect and intelligibility. HIDS was perceived to be more positive and intelligible than both FDS and ADS, with no difference between the latter two groups. These results show that people modify the acoustic properties of their voice in interactions with foreigners and hearing-impaired people, but that the benefits of these modification are more noticeable in HIDS than FDS.

1. Introduction

The relationship between hearing loss and the access that an individual with hearing loss has to spoken language is very complex [1]. The type of hearing loss an individual has is normally categorised into mild, moderate, moderately- severe, severe, severe-profound and profound [2]. Individuals at the mild end of the spectrum can normally achieve good access to spoken language when fitted with some type of hearing aid. However, those individuals at the other end of the spectrum who are severe-profound or profoundly deaf may suffer from an increased number of communicative problems. For example, Harkins and Bakke [2] found that children suffering from severe to profound hearing loss did not perceive speech in the same way as others without a hearing impairment even with amplification. The communication problems associated with hearing-impaired individuals and the steps that can be taken to improve communication (e.g. hearing aids & early intervention) have been previously investigated [3]. However, the way people communicate with hearing-impaired people is less well understood.

Previous research has examined speech directed to different audiences in order to investigate the prosodic modifications that individuals make to try to meet the communicative and emotional needs of different listeners [4]. This research has shown that people modify the acoustic properties of their voice to deal with the needs of their audience (e.g., [5]), and that such modifications benefit the listener (e.g., [5, 6]). For instance, clarity-enhancing modifications have been found in infant-directed speech [7], human-computer interactions [8], in speech directed to foreigners- [9] and hearing-impaired people (HIDS; [10]). However, studies investigating HIDS have used contrived scenarios, sometimes focusing on the speech of one person *imagining* talking to a hearing-impaired person [10]. Some studies have also included specific instructions (e.g., clear versus conversational speech, [11]). Recent studies (e.g., [12]) have shown that results from such studies are inconsistent with those from natural interactions. It is thus not clear how people actually modify their voice in interactions with hearing-impaired people, particularly if not specifically instructed to produce clear speech.

The present study aims to investigate whether people in interactions with hearing-impaired people a) modify their speech of their own accord without clear speech instructions, and b) how the listener perceives such modifications. We also investigate FDS to provide a linguistic comparison group to HIDS. Clarity-enhancing acoustic modifications have previously been found in FDS [9], and for this reason FDS may provide an excellent comparison group to HIDS. In line with previous research [7, 8, 9], we focus the acoustic analyses of the speech samples on intensity, mean F_0, F_0 range, mean duration and vowel hyperarticulation[1]. It has been suggested that acoustic modifications may result in listeners rating the speech to contain less positive affect [9]. Furthermore, the benefits of acoustic modifications for speech clarity have also been previously demonstrated [5, 6]. To explore this further, we also investigate how HIDS, FDS and ADS are perceived by other listeners with regards to clarity and positive vocal affect.

2. Method

2.1. Design

This study used a one way within-subjects experimental design. The independent variables were the types of speech recipients, consisting of: HIDS (condition 1), FDS (condition 2) and speech directed to an adult native speaker without a hearing impairment (ADS; condition 3). The study comprised two different components; a production and a perception component. For the production component the dependent

[1] Hyperarticulation here refers to vowel space expansion due to shifts in formant 1 (F1) and formant 2 (F2) of the three corner vowels /a/, /i/ and /u/.

variables were five different types of acoustic measures. These included mean intensity (measured in decibels), mean vowel duration (measured in seconds), vowel hyperarticulation (mean vowel space, measured in Hertz[2]), mean fundamental frequency (F_0) and F_0 range (measured in Hertz). For the perception component the dependent variables were ratings in relation to clarity and positive vocal affect (i.e., how positive the speech sample sounded).

2.2. DiapixUK task

The current study utilised the DiapixUK [14] task to obtain semi-standardised speech and keywords across the different conditions. This task has been recently developed to allow recording of natural speech across a variety of different speech conditions. The task usually involves two participants who take part in a 'spot the difference' task based on a cartoon image (Figure 1). The task was adapted so that participants would mention relevant keywords that would contain the three corner vowel /a/, /i/ and /u/. For instance, the cartoon depicted a shark and a sheep. Three different cartoons (farm, beach and street) were used to provide a different cartoon for each speech recipient group. Order of cartoons was randomised and counterbalanced.

Figure 1: *Modified version of DiapixUK [14] task.*

2.3. Production procedure

Ten Scottish native speaking female undergraduate psychology students (mean age 22.9, *sd* 3.84) from the greater Glasgow region were recruited through convenience sampling to serve as participants in the production component of this study. The participants had no known speech or hearing impairment at the time of the study. The participants were recorded whilst interacting with[2]: (1) an adult native speaking confederate (Scottish female, aged 21), (2) a foreign adult confederate, (1 of 3 females, either of Estonian nationality aged 21 or Czech nationality aged 21 and 22), and (3) a hearing-impaired adult confederate (Scottish female, aged 27). The foreign confederates had been living in Scotland for 2 ½ years and had noticeable accents. The hearing-impaired confederate is profoundly deaf and has been deaf since the age of 7[3]. She uses two hearing aids and mostly relies on lip reading.

2.4. Perception procedure

A further twenty British undergraduate psychology students recruited through convenience sampling were asked to participate in the perception component of this study. Participants consisted of five males and fifteen females (mean age 22.4, *sd* 2.14). The participants were asked to listen to 30 speech samples and then rate the samples using a Likert scale measuring positive vocal affect and clarity (adapted from [7 and 9]). None of the participants reported having a hearing impairment before taking part in the study.

2.5. Acoustic analyses

Analyses were carried out in Praat [15]. Consistent with previous research [9], the acoustic analyses focused on the three corner vowels (/a/, /i/ and /u/). All keywords containing the corner vowels that were uttered in the interactions were used in the analysis. Overall 362 words were analysed for mean vowel duration (seconds), mean F1/F2 values (Hz^2), mean intensity (dB), mean F_0 (Hz) and F_0 range (Hz). Mean F1/F2 values were used as x-y co-ordinates to plot bivariate vowel triangles for /a/, /i/ and /u/, from which vowel triangle area was calculated to detect 'hyperarticulation' of the vowels. Plotting and calculation of the vowel space was carried out in Autocad 2000.

3. Results

3.1. Production component

Five analyses of variance (ANOVA) were conducted to test the effect of the three speech recipient groups on the five acoustic measures[4]. Significant effects were followed up with pairwise comparisons (Bonferroni-adjusted).

There was no significant difference between the groups with regards to mean intensity, mean duration or mean F_0 range. There was a significant difference between the groups with regards to mean F_0 (F (2, 18) = 11.00, p = .001; partial eta squared = .55) and hyperarticulation of vowels (F (2, 18) = 6.41, p = .02; partial eta squared = .42). Participants produced significantly higher mean F_0 in HIDS and FDS than ADS (p = .005; p = .018 respectively). There was no significant difference in F_0 between HIDS and FDS (Figure 2). Participants also articulated their vowels more clearly in HIDS compared to ADS (p = .042). Interestingly, there was no significant difference in vowel hyperarticulation between FDS and the other two conditions (Figure 3), although the difference between FDS and ADS approached significance (p = .09).

[2] Speakers were informed that their interaction partners were deaf and foreign.

[3] Profound deafness results in sound detection of >90dB hearing level, and individuals with this type of hearing-impairment still benefit from hearing aids [i.e., 2].

[4] Low ratio between participant numbers and dependent variables did not permit multivariate analyses of variance.

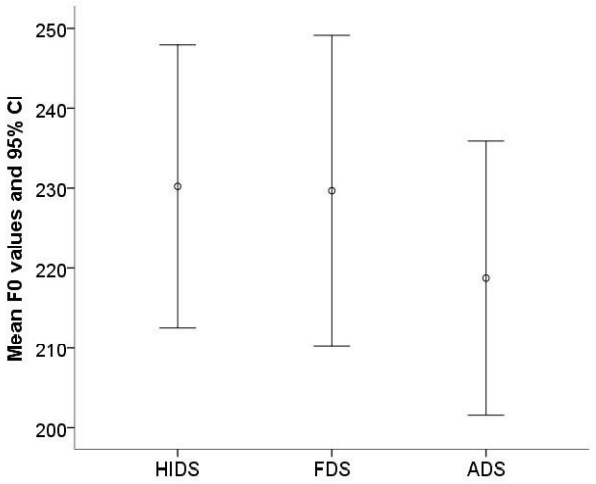

Figure 2: *Mean F_0 values (Hz) and 95% confidence intervals across the three speech recipient groups.*

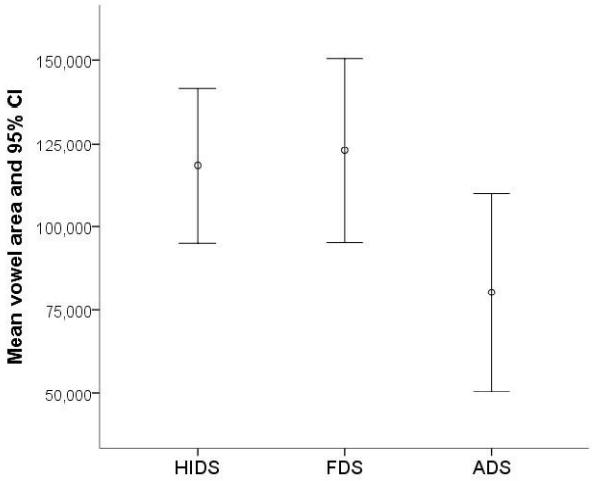

Figure 3: *Mean vowel triangle areas (Hz2) and 95% confidence intervals across the three speech recipient groups.*

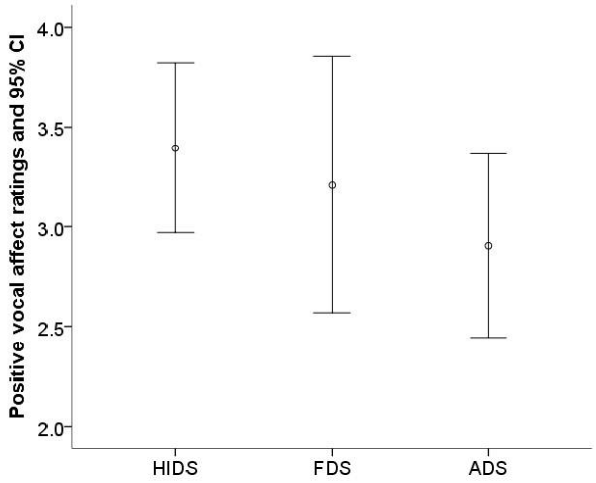

Figure 4: *Mean ratings for positive vocal affect and 95% confidence intervals across the speech recipient groups.*

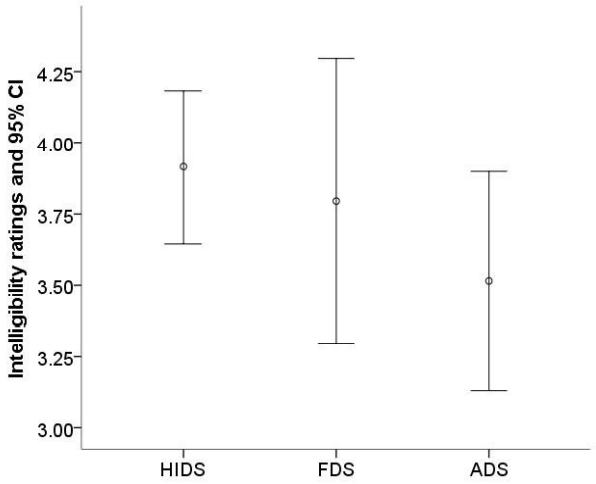

Figure 5: *Mean ratings for intelligibility and 95% confidence intervals across the three speech recipient groups.*

3.2. Perception component

Two ANOVAs were run to test the effects of the speech recipient groups on the two perceptual measures (positive vocal affect and intelligibility).

There was a significant difference between the three groups with regards to positive vocal affect (F (2, 18) = 7.27, p = .01; partial eta squared = .45) and intelligibility (F (2, 18) = 5.00, p = .02; partial eta squared = .36). HIDS was perceived to be significantly more positive than ADS (p =.002). No significant difference was found between ADS and FDS, or HIDS and FDS (Figure 4). HIDS was also rated as significantly more intelligible than ADS (p = .027) but not FDS. No difference was found between ADS and FDS (Figure 5).

4. Discussion

The aim of the current study was to investigate acoustic modifications in speech directed to people with a hearing-impairment and foreigners. The results of our study indicate that people adapt their speech modifications to suit the needs of their audiences. Participants in the current study employed a clarity-enhancing speech register in the interaction with the hearing-impaired person, as they articulated their vowels more clearly and increased their mean F_0. They also increased their mean F_0 in the interaction with the foreign confederates.

The findings with regards to HIDS mean F_0 are in line with previous research that has investigated the differences between HIDS and ADS using imaginary scenarios. Speakers seem to increase their mean F_0 in interactions with hearing-impaired listeners. Interestingly, in contrast to previous research [9] speakers in the present study also increased their F_0 in the interactions with the foreign listener. Our findings with regards to mean F_0 might explain why the raters perceived HIDS to contain more positive vocal affect than ADS. Increased F_0 has been associated with emotional affect [9], and the raters in the current study may have used the mean F_0 modifications for their rating decisions. However, in this case we would have expected FDS also to be perceived to contain significantly more positive vocal affect than ADS, which was not the case. Our findings here suggest a linguistic (clarity-enhancing) function for mean F_0, at least in adult interactions.

With regards to hyperarticulation, our findings are consistent with previous research investigating differences between clearer and conversational speech [5]. Speakers produced clearer vowels in the interactions with the hearing-impaired listener compared to the native adult listener without hearing impairment. Hyperarticulation of vowels occurred independent of vowel duration, which is line with Uther et al.'s [9] findings. This suggests that speakers adapt the acoustic modifications of their voice to support the communicative needs of their listener

Surprisingly, we did not find a significant difference in hyperarticulation between FDS and HIDS or FDS and ADS, although the latter approached significance. This means that articulation in FDS was more variable than articulation in the other two speech recipient groups, with some speakers adopting hyperarticulation while others did not. Our findings here are in contrast to previous research that found that participants hyperarticulated their vowels in interactions with foreigners compared to other native adult speakers [9]. One reason for this could be our choice of foreign confederates. Uther et al. [9] used a Chinese confederate in their interactions who had only been living in England for a couple of months. In the present study, we used three European confederates who had been living in Scotland for the last 2 ½ years and may have been more proficient in the English language. For this reason, speakers may not have hyperarticulated their vowels in these interactions because they felt that their foreign listener was capable of understanding their speech without this modification. Support for this comes from the perceptual component. Whilst HIDS speech was perceived to be more intelligible (i.e., articulated) than ADS, FDS was not perceived to be more intelligible than ADS. This suggests that speakers did not use clearer speech in FDS compared to ADS.

5. Conclusion

The results of the present study show that people adjust the acoustic properties of their voice to suit the communicative needs of their audience. Speakers adopted a specific speech register in their interaction with the hearing-impaired listener with subtle differences to their interaction with the foreign listener. These modifications resulted in a more positive and clearer perception by other listeners. Based on these findings, we suggest that there is a hierarchy of speech modifications, and that people adjust their speech register based on the perceived communicative needs of their audience.

6. References

[1] P. Kushalnagar, T.D. Topolski, B. Schick, T.C. Edwards, A.M. Skalicky and D.L. Patrick, "Mode of communication, perceived level of understanding, and perceived quality of life in youth who are deaf or hard of hearing," *Journal of deaf studies and deaf education,*16(4), pp. 512-523, 2011.

[2] J.E. Harkins and M. Bakke, "Technologies for communication: Status and trends," in M. Marschark & P. E. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education* (2nd ed., Vol. 1, pp. 425– 438). New York: Oxford University Press, 2011.

[3] A. R. Lederberg and J. Beal-Alvarez, "Expressing meaning: From communicative intent to building vocabulary," in M. Marschark and P. E. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education* (2nd ed., Vol. 1, pp. 258 – 275). New York, NY: Oxford University Press, 2011.

[4] Y. T. Matsuda, K. Ueno, K. Cheng, Y. Konishi, R. Mazuka and K. Okanoya, "Auditory observation of infant-directed speech by mothers: experiencedependent interaction between language and emotion in the basal ganglia," *Frontiers in human neuroscience*, 8, 2014.

[5] T. Bent, T and A.R. Bradlow, "The interlanguage speech intelligibility benefit," *Journal of the Acoustical Society of America,* 114(3), pp. 1600-1610, 2003.

[6] H.M. Liu, P. K. Kuhl, and F.M. Tsao, "An association between mothers' speech clarity and infants' speech discrimination skills," *Developmental Science,* 6, pp. F1-F10, 2003.

[7] D. Burnham, C. Kitamura, and U. Vollmer-Conna, "What's new, Pussycat? On Talking to Babies and Animals," *Sccience,* pp. 296, 1435, 2002.

[8] S. Oviatt, D. Levow, M. MacEachern and K. Kuhn, "Modelling hyperarticulated speech during human-computer error resolution," *Speech Communication,* 24(2), pp. 87-110, 1998.

[9] M. Uther, M.A. Knoll and D. Burnham, "Do you Speak E-NG-L-I-SH? A comparison of Foreigner and Infant Directed Speech," *Speech Communication,* 49, pp. 2-7, 2007.

[10] M.A. Picheny, N.I. Durlach and L.D. Braida, "Speaking clearly for the hard of hearing II: acoustic Characteristics of Clear and Conversational Speech," *Journal of Speech and Hearing Research,* 29, pp. 434-446, 1986.

[11] R. Smiljanić and A.R. Bradlow, "Production and perception of clear speech in Croatian and English," *Journal of the Acoustical Society of America,* 118 (3), pp. 1677-1688, 2005.

[12] M. A. Knoll, L. Scharrer and A. Costall, "Look at the shark": Evaluation of student- and actress produced standardised sentences of infant- and foreigner-directed speech," *Speech Communication*, 53, pp. 12-22, 2011.

[13] T. Baer, B.C. Moore and K. Kluk, "Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies," *Journal of the Acoustical Society of America,* 112, pp. 1133-1144, 2002.

[14] R. Baker, and V. Hazan, V. (2011). "DiapixUK: Task Material for the Elicitation of Multiple Spontaneous Speech Dialogs," *Behavior Research Methods,* 43, pp. 761-770, 2011.

[15] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [Computer program]. Version 5.4.08, retrieved 2014 from http://www.praat.org/.