ACOUSTICS 2012

# A-M I S-P-EA-K-I-NG C- L- E- AR- L- Y E-N-OU-GH?: An investigation of the possible role of vowel hyperarticulation in speech communication

J. Kangatharan, M. Uther, L. K. Kuhn and F. Gobet

Brunel University, Kingston Lane, UB8 3PH Uxbridge, UK
jayanthiny.kangatharan@brunel.ac.uk

Speech aimed at infants and foreigners has been reported to include the physical exaggeration of vowels, that is vowel hyperarticulation. Although infants have been demonstrated to experience hyperarticulated vowels in speech directed at them, little research has been done on whether vowel hyperarticulation occurs as a result of foreign appearance, foreign accent or as a consequence of both looking and sounding foreign. The present study therefore explored whether appearance and speech separately affect native speakers' hyperarticulation. Forty White British adult speakers communicated with one of four different confederate groups (2 types of appearance x 2 types of accent) to solve three modified 'DiapixUK' tasks. Results showed that vowel space across 4 vowels, including a diphthong did not significantly differ across experimental conditions. However, the interlocutor's physical appearance was observed to influence the second formant across target vowels indicating possible effects on speech intelligibility. Additional samples are being analysed to verify the results of the present study.

# 1 Introduction

Acoustic hyperarticulation of phonemes is the physical exaggeration of speech sounds [22]. One type of exaggeration is vowel hyperarticulation, which has been reported to occur in infant-directed speech (IDS) [1]. Vowel hyperarticulation also co-occurs with other acoustic changes in the speech register, including raised mean *F*0 (pitch) and exaggeration in pitch contours. Positive emotional affect is also rated highly in IDS compared to adult directed speech (ADS) [1, 2, 13, 17, 20].

Research has suggested that vowel hyperarticulated speech might help infants to learn phonetic units [12, 26]. For example, the vowels /i/, /a/ and /u/ were reported to be acoustically more exaggerated in IDS than in ADS [12]. Specifically, it was observed that the spacing between the vowels (measured by area in *F*1 and *F*2 space) is larger in IDS compared to ADS [12]. This hyperarticulation has been suggested to have didactic utility for the infant in acquiring new phonetic categories [12]. Accordingly, native-like phonetic categorization has been demonstrated in infants who were exposed to good examples of phonetic categories [7, 9, 11]. Similarly, adults choose hyperarticulated vowels as improved examples of vowel categories [10].

## 1.1 Vowel hyperarticulation in speech to adults

In ADS, phonetic units are frequently not very well stipulated as a result of which different phonetic categories cannot be specified [8]. Thus, the phonetic units used in ADS might lead to a poor linguistic learning outcome for infants since information that is required to learn language is hypospecified in ADS [6]. The idea that hyperarticulation is a uniquely didactic acoustic feature was tested in a study by Burnham et al. [5] who showed that pet-directed speech contained pitch and affect patterns similar to infant-directed speech but lacked any hyperarticulation effects. Adult language learners, who communicate with adult speakers in a second language (L2), may also therefore require similar acoustic modifications (i.e. hyperarticulation) as those given to infants so that they can learn the new categories. This idea was tested in a study in which participants spoke to confederates that were native speakers or who were foreigners (both looked and sounded non-native). It was found that the foreigner-directed speech (FDS) had vowels that were hyperarticulated compared to those that were normal adult-directed (native confederate) [21].

## 1.2 Vowel hyperarticulation in speech to individuals that appear or sound 'foreign'

Despite the presumed utility of hyperarticulation as a didactic tool for learning new phonemes, this idea has not been directly empirically tested. Numerous studies on the intelligibility of different types of speech have reported the use of hyperarticulated vowels in IDS [3, 8, 15]. Liu et al. [15], for example, analyzed the benefit of mothers' more comprehensible speech for infants. Although they showed that mothers' more comprehensible speech has an advantage for infants' speech perception, only a correlation between mothers' vowel hyperarticulation and their infants' speech perception capabilities could be ascertained.

Moreover, little is known about whether native speakers' speech at other native speakers is different from their speech at native looking and foreign sounding individuals, on the one hand, and at foreign looking and native sounding individuals, on the other hand. Past research on different speech registers used with foreigners has focused on those with both a foreign appearance and foreign accent [14, 16] (**f**oreign **l**ooking and **f**oreign **s**ounding: **FLFS**).

To this end, this study sought to separate the variables of looking and sounding foreign. In this study, there are four kinds of interlocutors: those who both look and sound native (**n**ative **l**ooking and **n**ative **s**ounding: **NLNS**), those who appear foreign but linguistically sound like native speakers (**f**oreign **l**ooking and **n**ative **s**ounding: **FLNS**), those who appear native but sound foreign (**n**ative **l**ooking and **f**oreign **s**ounding: **NLFS**) and those who both look and sound foreign (**f**oreign **l**ooking and **f**oreign **s**ounding: **FLFS**). There is a strong theoretical and practical importance in investigating these variables separately so that it can be ascertained whether speech registers are most affected by either the appearance or accent of the interlocutor.

## 1.3 Aim of the present study

This study investigated whether the physical appearance or accent of the interlocutor results in independent effects on eliciting hyperarticulation of a native speaker. This study specifically looked at the vowels /ai/ (as in 'sign'), /e/ (as in 'yellow'), /iː/ (as in 'green') and /ɔː/ (as in 'door'). It was hypothesized that vowel hyperarticulation would occur more in speech to FLFS speakers than in speech to speakers who are FLNS, NLFS and NLNS. It was also predicted that speech to NLFS

speakers would be more hyperarticulated compared to speech to FLNS and NLNS speakers.

## 2 Methods

### 2.1 Speakers and confederates

Forty White British speakers between 18 and 35 years were asked to communicate with one of four different speaker groups (10 White British individuals, 10 speakers of White European ethnicity with native White British appearance and foreign accent, 10 speakers of Asian (Indian/Pakistani or Bengali) ethnicity with foreign appearance and native accent, and 10 speakers of Asian ethnicity with foreign appearance and foreign accent). Participants were recruited from the student population of Brunel University.

### 2.2 Design

This study used a 2 (confederate's accent: native, foreign) x 2 (confederate's physical appearance: native, foreign) between-subjects design. Thus, the two independent variables formed an experimental design with four groups of different confederate types: NLNS, NLFS, FLNS and FLFS. The dependent variable was the degree of hyperarticulation in the target words in which one of the four target vowels was present.

### 2.3 Materials

For the purpose of eliciting the target vowels /ai/, /e/, /i:/ and /ɔ:/, the words 'sign', 'yellow', 'green' and 'door' were chosen as specific target words, which contain each of the target vowels respectively. To facilitate the elicitation of these target vowels from the native speakers, three "Spot-the-difference" (Diapix) tasks were used. These tasks were modified versions of the tasks developed by Baker and Hazan [2]. The first picture depicted a beach scene, the second a farm scene and the third picture a street scene. A digital voice recorder Edirol R-09HR by Roland (sampling rate: 44.1 kHz) was used to record all verbal interactions. Each interaction was recorded as a mono 16-bit wavfile.

### 2.4 Procedure

In each half an hour audio-recorded interaction, two participants were seated opposite each other and each participant received a folder with three pictures, each illustrating a different scene. For each scene, there were 13 differences between the picture that one participant received and the picture of her partner. The differences included an absent object or an alteration to one of the objects on the picture. Participants were instructed to work together to verbally find out the differences between their pictures. The task lasted about ten minutes and was followed for all three pictures. Adult native speakers were then asked to read out short lists of sentences which contained target words common to those elicited in the Diapix task (as a comparison between natural and read speech). Demographic and linguistic background information was collected from each participant.

## 3 Results

Phonetic (vowel quadrilateral area) measures were analyzed using Multivariate Analysis of Variance (MANOVA). Formant values for *F*1 and *F*2 in the vowels of the target words 'sign', 'yellow', 'green' and 'door' were calculated with the software application Praat 5.2.27. It is noteworthy that the extent of the differences between the target vowels was not as marked as they have been in past research. Nonetheless, this result is comprehensible if one considers the use of different corner vowels in this study and the dissimilarities between the South East English dialect in comparison to the Australian and American English dialects in past research.

Three cases from the NLNS condition, two cases from the NLFS condition and one case from the FLNS and from the FLFS conditions were removed as outliers before a MANOVA was performed. The analysis showed no significant main effect of speech ($F(9, 21) = .591$; $p > 0.05$). However, appearance was revealed to be a significant main effect ($F(9, 21) = 2.37$, $p = 0.05$; partial eta squared = .504, a 'large' effect). There was no significant interaction (Speech x Appearance: $F(9, 21) = .459$, $p > 0.05$). Tests of between-subjects effects demonstrated that the type of appearance had a significant effect on the second formant for the vowel /ai/ ($F(1, 29) = 6.34$, $p < 0.05$; partial eta squared = .18, a 'large' effect) and for the vowel /e/ ($F(9, 21) = 2.37$; $p < 0.05$; partial eta squared = .18, a 'large' effect).

The formant values for *F*1 and *F*2 in the vowel of the target words 'sign', 'yellow', 'green' and 'door' including the resulting vowel quadrilateral area shown in Figure 1.

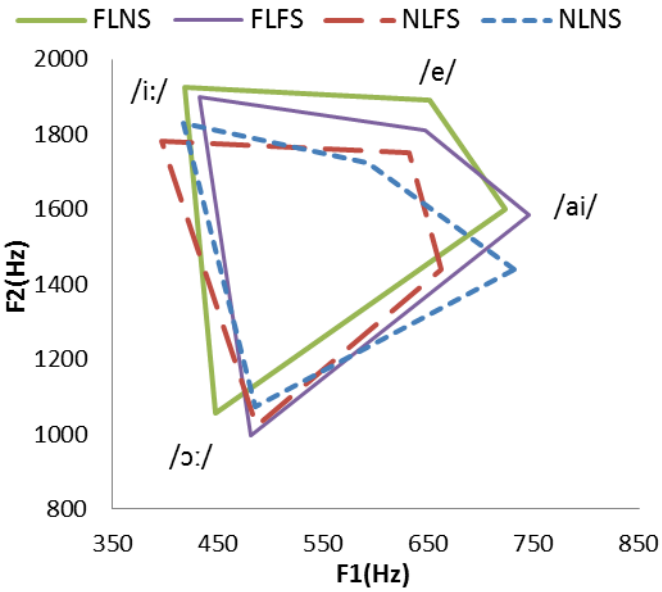


Figure 1: Vowel areas between the four target vowels in the four conditions (NLNS, NLS, FLNS and FLFS) as indicated by differences in *F*1 and *F*2.

An analysis of mean quadrilateral areas showed no significant differences in vowel space across all four conditions. A repeated measures analysis of variance (ANOVA) was executed to further investigate the effect of appearance on *F*1 and *F*2. In addition to confirming the observed result that appearance influences the second formant but not the first, the analysis showed that the

significant main effect of appearance reflected another significant difference which is that *F*2 is significantly larger for the foreign looking conditions than for the native looking conditions ($F(1, 29) = 9.18$, $p < 0.05$; partial eta squared = .24, a 'large' effect). Tests of within-subjects contrasts revealed a significant interaction between vowels and appearance for the second formant, indicating that the mean values for *F*2 found in the foreign looking conditions are different across vowels ($F(1, 29) = 4.31$, $p < 0.05$; partial eta squared = .13, a 'medium' effect).

# 4 Discussion

Contrary to the hypotheses that we set out to test, the results of this study indicate no differences in the manner in which White British native speakers communicated with native sounding speakers compared to their way in which they addressed foreign sounding individuals. This is in contrast to findings from previous studies in which speech towards foreign sounding individuals was vowel hyperarticulated compared to speech directed at native speakers [21]. One reason that might have contributed to this result is the fact that this study used a between-subjects design while past research employed the same White British speakers across all conditions [21]. As a consequence, the absence of a significant main effect of speech in this study may have been influenced by individual differences among speakers in the 40 speech samples.

Apart from inter-subject variability, a further aspect that might have influenced this result is the degree to which White British native speakers were familiar with the ethnic group that the speakers (who vary in accent and appearance) belong to. Demographic details revealed for example that 44.4% of White British native speakers in the FLFS condition encounter members of the FLFS group many times a day and also are befriended with them. Similarly, 55.5% of White British native speakers in the FLNS condition meet members of the FLNS group many times a day and are friends with them as well. In both the FLNS and FLFS condition, 66.67% of the native speakers expressed their familiarity with members of their dialogue partner's ethnic group. In contrast, 50% of White British native speakers in the NLFS condition said to be not very familiar with members of that group. They said they meet NLFS individuals on a monthly basis while 12.5% of native speakers stated that they would meet NLFS individuals many times a day. Thus, it can be said that another reason for the present results is the fact that some native speakers knew their dialogue partners prior their participation in the study, which likely would have contributed to increased variability in the data.

The lack of significant main effect of speech in this study can probably also be attributed to the good proficiency of NLFS, FLNS and FLFS speakers in the English language. Accordingly, although all FLFS speakers use their mother tongue (L1) to communicate with the people they talk to most in their lives compared to 62.5% of NLFS speakers and 33.3% of FLNS speakers, all speakers stated to be able to easily take part in complex discussions and to be confident in reading more complex material such as long novels or academic writing. This observation is supported by the White British native speakers' statement that the majority of their dialogue partners neither slowed them down in task completion nor could have done more to solve the task. Although White British native speakers were able to identify the accents of their dialogue partners in the NLFS and FLFS conditions as foreign, all native speakers found their dialogue partners to be helpful and friendly across all conditions. Compared to 88.89% of native speakers in the FLFS condition, all native speakers in the other conditions felt it was easy to communicate with their dialogue partners.

A further reason for the lack of significant findings with regard to speech is the fact that a diphthong was included in the phonetic analysis. The inclusion of the diphthong '/ai/' may have influenced the results for the vowel space area. This is because in contrast to 'pure' vowels that have only one vowel sound, diphthongs are vowels of changing resonance in which two adjoining vowel sounds transpire in one syllable. Thus, it can be said that the lack of significant differences in the native speakers' speech to individuals in the four different conditions might have been caused by inter-subject variability, native speakers' increased familiarity with members of the ethnic group to which the NLFS, FLNS and FLFS speakers belong and the NLFS and FLFS speakers' high levels of proficiency in L2 as well as the inclusion of a diphthong in the analysis. Moreover, contrary to previous studies, the present study was conducted in a speech cubicle which represents an artificial environment [21]. This aspect additionally might have contributed to an absence of speech that includes the exaggeration of vowels toward foreign looking interlocutors.

Nonetheless, this study provides strong evidence for the important role of the second formant in speech intelligibility. Past research identified the second formant to be critical in speech intelligibility [18, 19]. In line with these previous studies, the present study has shown that the perception of foreign physical appearance exerted an effect on the second formants across all vowels. Specifically, it can be said that 24% of the mean values for the second formant for each vowel can be accounted for by the foreignness of the interlocutor's physical appearance. The results of this study can therefore be suggested to experimentally support previous research by showing that in contrast to the first formant, the second formant is an essential contributor to speech intelligibility [18, 19].

# 5 Conclusion

Thus, although, the present study seems to suggest that physical appearance and not speech plays a significant part in speech communication, additional samples are currently being analysed for each of the four experimental conditions with the omission of the diphthong and also examining other vowels to find out whether the absence of statistical differences across conditions might be attributed to the choice of vowel. Moreover, acoustic (*F*0) and affective measures from the present samples are also being analysed to discover how the resulting outcomes will change for the different experimental conditions.

## References


[1] J. E. Andruski, P. K. Kuhl. "The Acoustic Structure of Vowels in Mothers' Speech to Infants and Adults", *Proc. 4th ICSLP* Philadelphia, 1545-1548 (1996)

[2] R. Baker, V. Hazan, "Acoustic-phonetic characteristics of naturally-elicited clear speech in British English" *157th ASA Meeting* Oregon (2009)

[3] N. Bernstein Ratner, "Patterns of vowel modification contained by mother-child speech" *Journal of Child Language*, 11, 557-578 (1984)

[5] D. Burnham, C. Kitamura, U. Vollmer-Conna," What's New, Pussycat? On talking to Babies and Animals*", Science*, 296, 1435 (2002)

[6] N. Chomsky, Rules and Representations. Columbia University Press: New York (1981)

[7] D. Grieser, P. K. Kuhl,"Categorization of speech by infants: Support for speech-sound prototypes", Developmental Psychology, 25, 577-588 (1989)

[8] J. Hillenbrand, L. A. Getty, M. J. Clark, K. Wheeler, "Acoustic characteristics of American English Vowels" *J. Acoust. Soc. Am*., 97, 3099-3111 (1995)

[9] P. Iverson, P. K. Kuhl, "Mapping the perceptual magnet effect for speech using detection theory and multidimensional scaling", *J. Acoust. Soc. Am*., 97, 553-562 (1995)

[10] K. Johnson, E. Flemming, R. Wright, "The hyperspace effect: phonetic targets are hyperarticulated" *Language*, 69, 505-528 (1993)

[11] P. K. Kuhl, K. A. Williams, F. Lacerda, K. N. Stevens, B. Lindblom, Language experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606-60 (1992)

[12] P. K. Kuhl, J. E. Andruski, I. A. Chistovich, L. A. Chistovich, E. V. Kozhevnikova, V. L. Ryskina, E. I. Stolyarova, U. Sundberg, F. Lacerda, "Cross-Language Analysis of phonetic Units in Language Addressed to Infants", *Science*, 277, 684 (1997)

[13] P. K. Kuhl, " Early language acquisition: cracking the speech code", *Nature Reviews Neuroscience*, 5, 831-842 (2004)

[14] L. N. Littleford, M. O.- D. Wright, M. Sayoc-Parial,"White Students' Intergroup Anxiety During Same- Race and Interracial Interactions: A Multimethod Approach. Basic and Applied" *Social Psychology*, 27, 85-94 (2005)

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

[16] C. E.Snow, R. van Eeden, P. Muysken, "The interactional origins of foreigner talk: Municipal employees and foreign workers", *International Journal of the Sociology of Language*, 81, 81-91 (1981)

[17] H. Kim, M. M. Diehl, R. Panneton, C. Moon, "Hyperarticulation in Mothers' Speech to Babies and Puppies" *XVth Biennial International Conference on Infant Studies*, Japan (2006)

[18] I. B.Thomas, "The influence of first and second formant on the intelligibility of clipped speech", *Journal of Audio Engineering Society*, 16, 182 (1968)

[19] I. B.Thomas, R. J. Niederjohn "The intelligibility of filtered-clipped speech in noise", *Journal of Audio Engineering Society*, 16, 182 (1968)

[20] L. J. Trainor, R. N. Desjardins, "Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels", *Psychonomic Bulletin and Review*, 9, 335-340 (2002)

[21] M. Uther, M. A. Knoll, D. Burnham, " Do you speak E-n-g-l-i-s-h? A comparison of foreigner- and infant-directed speech" *Speech Communication*, 49, 2-7 (2007)

[22] D. H. Whalen, H. S. Magen, M. Pouplier, A. Min Kang, K. Iskarous, "Vowel Production and Perception: Hyperarticulation without a Hyperspace Effect", *Language and Speech*, 47, 155 (2004)