Infant Directed Speech expands F1/F2 space for /i a u/: But what happens to central vowels /ɜ ʊ/ in Australian English?

Antonia Goetz[1], Eylem Altuntas[1], Nicole Swinfield[1], Denis Burnham[1], & Catherine T. Best[1]

[1]MARCS Institute for Brain, Behaviour & Development, Western Sydney University, Sydney, Australia
a.goetz@westernsydney.edu.au

ABSTRACT

Relative to their adult-directed speech (ADS), mothers expand the triangle circumscribed by the reference vowels (/i a u/) in F1/F2 space when interacting with their infants. This vowel space expansion in infant-directed speech (IDS) is thought to support word learning and phonological development. But what about non-reference (i.e., 'central') vowels such as /ɜ ʊ/? Are they also located differently in F1/F2 space in IDS and ADS? We compared /ɜ ʊ/ in Australian English, in which they fall posterior to the vowel triangle, produced by 15 mothers addressing their 4-month-old versus an adult. We extracted F1 and F2 values and calculated the Euclidean distance between the vowels for IDS and ADS. The distance was significantly *shorter* in IDS than ADS, rather than longer as with the /i a u/ triangle. F2 (but not F1) was also lower in IDS than ADS, suggesting more posterior tongue placement for /ɜ ʊ/ in IDS.

Keywords: infant-directed speech, central vowels, speech production, vowel hyperarticulation

1. INTRODUCTION

When people talk to infants, they intuitively use a speech register called infant-directed speech (IDS), which is different from the speech used in communication with other adults (adult-directed speech, ADS). The acoustic features of IDS relative to ADS include higher average fundamental frequency (f0) [1], greater variation in f0 [1, 2] and vowel hyperarticulation [3], which refers to expansion of the triangle circumscribed by /i a u/ in F1/F2 acoustic space in IDS relative to ADS. The phonological reference vowels, /i a u/, mark the extremes of the two primary dimensions of vowel systems: height and anterior-posterior position of the tongue body. All languages with three or more vowels have /i a u/; numerous languages have only /i a u/.

The altered acoustic features of IDS are posited to serve multiple functions in communication with infants. The higher f0 of IDS, and its associated positive vocal affect, help to regulate infants' emotional states [4]; and /i a u/ hyperarticulation in IDS is thought to facilitate early lexical development [5]. For example, mothers' /i a u/ vowel hyperarticulation in communication with their 9-month-old infants is a reliable predictor of expressive vocabulary size at 19 months of age [3], and hyperarticulated /i a u/ are linked to facilitation of infants' word recognition also at 19 months [6]. While it has been hypothesised that caregivers hyperarticulate vowels in order to clarify the acoustic-phonetic features of the vowels [7], /i a u/ hyperarticulation has not always been found in IDS in languages other than English [8-10], and the underlying mechanism and functions of IDS remain unclear [11, 12] (but see, [13]). The specific unresolved issue our study addresses is how *non-reference*, i.e., *central*, vowels are produced in IDS versus ADS, and whether any differences results are associated with their phonological status as non-reference/central in the vowel system, or with their position in F1/F2 space. For this purpose, we investigated the production of the central vowels /ɜ ʊ/ in IDS and ADS in Australian English.

Previous studies with Australian English (AusE) have found that the /i a u/ vowel triangle defined by the vowels in *sheep* (/i/), *shark* (produced as /a/ in AusE, a non-rhotic accent) and *shoes* (/u/) is indeed expanded in IDS compared to ADS [3, 14]. However, in AusE /u/ is articulated in a much more fronted (more anterior) position (realised phonetically as [ʉ]), and /a/ is slightly more fronted, than they are in most other varieties of English (e.g., Received Pronunciation [RP] British English), as illustrated in Figure 1. That is, two of the three reference vowels, /u/ and /a/, are not articulated at the most extreme posterior positions possible in the acoustic vowel space, as they are in RP British English. Consequently, the AusE vowel triangle has become highly truncated in the anterior-posterior (F2) dimension, i.e., compressed into the front half of the oral portion of the vocal tract (see [15]). And in relation to that truncation of the reference vowel triangle in F1/F2 space, the central vowels /ɜ ʊ/ (the vowels in *shirt* and *chook*, respectively) actually fall outside rather than inside the triangle In acoustic space, as they do in RP British. So, these two *phonologically* central (non-reference) vowels are *acoustically* more posterior than AusE /u/ and /a/. This situation is more extreme for /ʊ/ than for /ɜ/ but even the latter is slightly posterior to the /u/-/a/ edge of the reference vowel triangle instead of being squarely within the triangle. The open question is: What happens to the phonologically central but acoustically peripheral vowels of AusE that are *outside* the vowel triangle? For instance, are these vowels also hyperarticulated, i.e., spread further apart as are the reference vowels) or hypoarticulated, i.e., more spatially congested in IDS compared to ADS, or are they simply unaffected by the change of register?

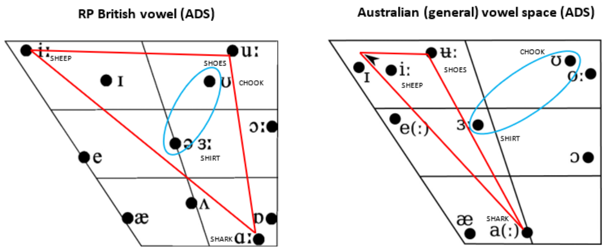

Figure 1: Vowel triangles (red) defined by /i a u/ in F1/F2 space and ‘central’ vowels /ɜ ʊ/ (turquoise) in British English (left, [16] and AusE (right, [17]. The vertical dimension is vowel height (F1, reverse scale – high F1 = low vowels); the horizontal dimension is anterior-posterior tongue position (F2, reverse scale – high F2 = front vowels).

As part of the [11] project, this study aimed to compare the acoustic features of two AusE vowels in *shirt* (/ɜ/) and *chook* (/ʊ/)[1] produced in IDS and ADS. As we have noted, in AusE these vowels fall outside of the traditional phonologically defined vowel triangle. This situation allows us to compare the acoustic features of a phonologically central vowel (/ʊ/) that is phonetically more extreme in F1/F2 space than the high back reference vowel /u/, to those of another phonologically central vowel (/ɜ/) that is closer to the centre of F1/F2 space but still not inside the reference vowel triangle (see Fig. 1).

Only one previous study has compared IDS and ADS for a vowel that falls *inside* in the vowel triangle in American English [11]. That study compared an anterior non-reference vowel, /ɪ/, to the closest reference vowel /i/, and found the distance between them to be shorter in IDS than in ADS, even though the reference vowels were hyperarticulated [12]. However, based on those results for American English we cannot draw conclusions about hyper- or hypoarticulation of the ‘central’ vowels in AusE, given that in AusE the phonologically ‘central’ vowels fall outside of the vowel triangle (Fig. 1).

2. METHOD

2.1. Participants

Fifteen mother-infant dyads participated in this study as part of a longitudinal project. All mothers were native speakers of AusE, and all infants (8 female) were acquiring only AusE as their first language with no more than 4 hours per week of exposure to any language other than English. No infants had any reported health issues, and none were at risk for any language or cognitive disabilities.

Infants participated in an auditory-visual speech perception experiment at the ages of 4, 7 and 10 months, and at each age IDS and ADS recordings were conducted after the perception experiment. Only the IDS session at 4 months (M_{age} = 143 days, range 3.95 – 5.62 months) is included in this report. However, as mothers produced less target words in ADS than IDS, the ADS recordings for all three ages were analysed here to increase the number of ADS productions of the target words in ADS (see *2.2. Procedure*). Infants received a diploma and toy. Caregivers gave informed consent.

2.2. Procedure

For the IDS and ADS sessions, mothers and infants were assigned to separate rooms connected via an audio-visual double screen set-up. This allowed communication between the mother (wearing a microphone-fitted headset (DPA) and infant (or experimenter, for the ADS sessions) in real time with the additional advantage that both participants’ visual expressions could be recorded. Only the audio recordings are reported here.

In the IDS session, mothers were provided with five toy objects to prompt elicitations of multiple productions of the target words *sheep*, *shoes*, *shark*, *shirt*, and *chook* (see labels in Fig. 1). Before the IDS sessions, mothers were presented with video-recorded instructions spoken by a native speaker of AusE to ensure that they were introduced to the target words with Australian-English pronunciation. The instructions included directions for the mothers to play naturally with their infant by using the five target-named objects, and to aim to use each word multiple times. Immediately following the IDS recording sessions, the experimenter interviewed the mother in the same double video set-up to elicit productions of the target words in communication with another adult. The IDS and ADS sessions lasted between 10 and 15 minutes in total.

2.3. Acoustic analysis

This report concerns the audio productions of the two central vowels (/ɜ/ and /ʊ/) as in AusE *shirt* and *chook*. As noted above, these vowels are of interest because they fall outside of the phonologically-defined vowel triangle in AusE. All audio files were annotated and acoustically analysed using PRAAT [19]. First, we annotated the onset and offset of the words and vowels manually using visual and auditory judgments. Segments that contained acoustic artefacts (e.g., vocalisation by the infant, or heavy glottalization) or sung speech to the infant were excluded from analyses. The mean number of extracted tokens per speaker for each vowel in each register (IDS, ADS) is shown in Table 1.

Table 1: Mean numbers per mother (*standard deviation*) of extracted target vowels used for the acoustic analysis of IDS and ADS.

Vowel	IDS	ADS
shirt /ɜ/	8.9 *(4.2)*	6.3 *(2.4)*
chook /ʊ/	13.9 *(7.3)*	7.4 *(1.9)*

Second, we used a PRAAT script to extract the values of the first formant (F1), second formant (F2), and third formant (F3) at 40% and 80% from the vowel onset to identify the steady-state nucleus. Third, we averaged all articulations of each vowel in each register per mother and calculated the Euclidean distance (*ed*) between the mean position of the /ɜ/ and /ʊ/ vowels in each register using the equation:

$$ed = \sqrt{(F1_{/ɜ/} - F1_{/ʊ/})^2 + (F2_{/ɜ/} - F2_{/ʊ/})^2}$$

Lastly, we calculated the Pillai score [19] to measure the overlap between two vowel cluster. The Pillai score value range between 0 (indicting almost perfect overlap) and 1 (no overlap).

3. RESULTS

Statistical analysis was performed using R [21] and lme4 [22], and plots were created by using ggplot2 [22]. Figure 2 shows the averaged F1 and F2 values

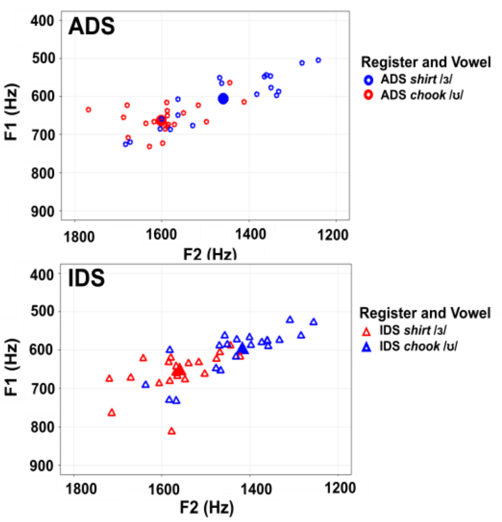

Figure 2: Mean F1 and F2 (Hz) of each mother for each Vowel in each Register (ADS above, IDS below). Filled markers indicate mean values for each vowel in each register.

per mother for each Vowel (*shirt* /ɜ/ and *chook* /ʊ/) in each speech Register (ADS, IDS).

We performed statistical comparisons with two mixed effect models to compare (i) the Euclidean distance, (ii) the Pillai score and (iii) the F1 and F2 values between the two vowels in ADS and in IDS.

For the first analysis, the dependent variable was Euclidean distance. Register (IDS coded as +0.5 vs. ADS coded at -0.5) was entered as a fixed effect and Mother as a random effect in the model. The results revealed that the Euclidean distance between the /ɜ/ and /ʊ/ vowels was significantly shorter in IDS than ADS (β *(SE)* = -83.26 (15.78), t = -5.276, p < .001), see Figure 3.

For the second set of analysis we extracted the Pillai score, by using the manova() function in R. The Pillai score was higher for ADS (Mean = 0.32, SD = 0.20) than IDS (Mean = 0.21, SD = 0.23).

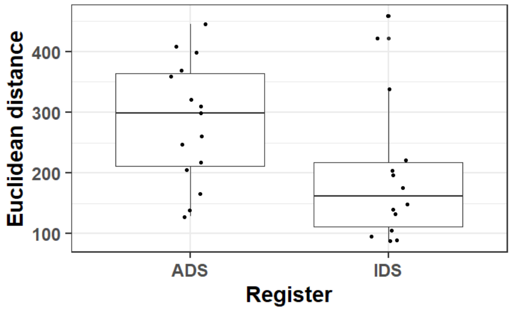

Figure 3: Boxplot of the Euclidean distances between the vowels *shirt* /ɜ/ and *chook* /ʊ/ in ADS and IDS. Dots represent the Euclidean distance per mother in each Register, averaged over tokens.

For the third analyses the dependent variables were F1 and F2 values in Hz with Register (IDS coded as +0.5 vs. ADS coded at -0.5) interacting with Vowel (*shirt* coded as 0.5, *chook* coded as -0.5) as fixed factors, and Mother as a random effect.

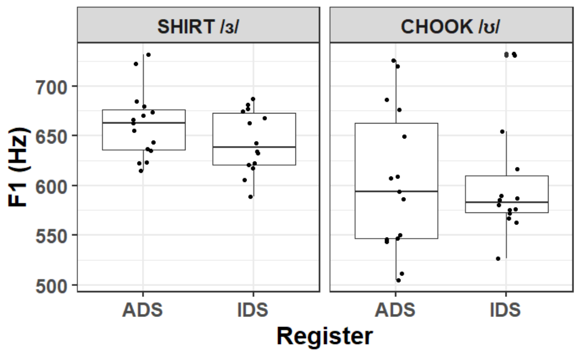

Figure 4: F1 (Hz) for the two Vowels (*shirt* /ɜ/ on the left and *chook* /ʊ/ on the right).

The F1 analysis (see Figure 4) showed neither a significant main effect of Register (β *(SE)* = -4.328 (6.370), t = -0.689, p = .500), nor a significant interaction with Vowel (β *(SE)* = -4.472 (6.338), t = -0.706, p =.484).For the F2 analysis, however, there was a main effect of Register; F2 was significantly lower in IDS than ADS (*β (SE)* = -28.363 (11.472), *t* = -2.472, *p* = .018, see Figure 5), but still showed no significant interaction with Vowel (*β (SE)* = -4.755 (11.392), *t* = -0.417, *p* = .679). These F2 findings indicate that there is a more posterior tongue placement for both vowels in IDS than ADS, i.e., both vowels are positioned farther beyond the reference vowel triangle in IDS than ADS.

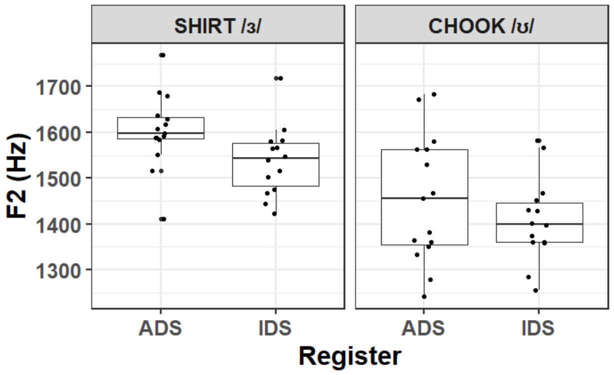

Figure 5: F2 (Hz) for the two Vowels (*shirt* /ɜ/ on the left and *chook* /ʊ/ on the right).

4. DISCUSSION

In contrast to documented expansion of the F1/F2 space circumscribed by the phonological reference vowels /i a u/ in IDS compared to ADS, little is known about whether mothers also change the acoustic features of non-reference vowels when speaking to their infants versus to an adult, and if so, then in what direction (expansion, contraction) such changes might occur. The present study investigated mothers' IDS versus ADS articulation of vowels that function phonologically as non-reference vowels, but that fall outside rather than inside the reference vowel triangle in their phonetic details, in AusE.

We found that Australian mothers produced the non-reference/central vowels /ɜ/ and /ʊ/ with a shorter Euclidean distance and with more overlap in IDS than in ADS, suggesting that these vowels are hypo- rather than hyper-articulated relative to each other in IDS. In addition, lower F2 values in both vowels in IDS than ADS indicate that these vowels are produced with a more posterior tongue placement in IDS than in ADS.

Previous studies with AusE mothers have shown that they reliably expand their vowel space in communication with their infants, via hyperarticulation of the /i a u/ reference vowels. However, results regarding hyperarticulation are not entirely consistent across languages. While some studies have found that reference vowels are hyperarticulated [3, 9], others have found that the reference vowels may be hypoarticulated in languages other than English [8-10] in IDS compared to ADS. Even in English hyperarticulation is also lacking under certain external conditions such as experimental manipulation of the infant's ability to hear their mother's voice [24], or with inherent factors such as infants' familial risk for dyslexia [14].

In this study we have shown that while the reference vowels, /i, a, u/, are hyperarticulated in AusE, the distance between two non-reference vowels is hypoarticulated. This is consistent with [12] which found that, even when the reference vowels are hyperarticulated, the distance between the non-reference central vowel /ɪ/ that falls *inside* the reference vowel triangle and the reference vowel /i/ was reduced in IDS, i.e., hyp*o*articulated, compared to ADS. The co-occurrence of *hyper*articulation of reference vowels and *hypo*articulation of a non-reference vowel could suggest that mothers apply both acoustic strategies to support infant linguistic development. Our results with two non-reference vowels that fall *outside* the reference vowel triangle suggest some additional insights. While these two non-reference vowels, which fall outside the /i a u/ F1/F2 triangle in AusE, are hypo-articulated in relation to each other, they are both articulated farther posterior in IDS than ADS. It is possible that the more posterior tongue placement increases the distance between these and the reference vowels, i.e., the difference between these non-reference vowels and the reference vowel triangle may, in contrast to [14] be *hyper*articulated in AusE. If so, that could facilitate infants' ability to differentiate between reference and non-reference vowels in this variety of English.

These observations raise unanswered questions as to (i) whether hypo- or hyperarticulation of non-reference vowels with respect reference vowels depends on their phonetic position inside or outside the reference vowel triangle, and (ii) whether IDS modifications of reference and non-reference vowels, have didactic value for infant language learning. So, future research is required: (i) to compare productions of /ɜ/ and /ʊ/ to those of /i a u/, as well as to productions of other non-referent vowels that fall *inside* the /i a u/ triangle in AusE (e.g., /ɪ/ as in *chip*); and (ii) to examine how the degree of reference vowel triangle expansion *and* modifications of non-reference vowels in a mother's speech to her infant at an early age (e.g., 4 months) may predict her infant's vocabulary [as per 4], and word recognition [as per 7] at 19 months of age. In addition, given that the two extant studies with non-referent vowels investigated different English varieties, IDS versus ADS articulations of the *shirt* and *chook* vowels should also be examined in other English varieties in which these vowels fall *inside* the reference vowel triangle (e.g., American, RP British).

5. REFERENCES

[1] R. P. Cooper and R. N. Aslin, "Developmental differences in infant attention to the spectral properties of infant-directed speech," *Child development*, vol. 65, no. 6, pp. 1663-1677, 1994.

[2] A. Fernald, "Intonation and communicative intent in mothers' speech to infants: is the melody the message?," *Child development*, pp. 1497-1510, 1989.

[3] M. Kalashnikova and D. Burnham, "Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions," *Journal of child language*, vol. 45, no. 5, pp. 1035-1053, 2018.

[4] M. Papoušek, "Communication in early infancy: An arena of intersubjective learning," *Infant Behavior and Development*, vol. 30, no. 2, pp. 258-266, 2007.

[5] P. K. Kuhl, "Early language acquisition: cracking the speech code," *Nature reviews neuroscience*, vol. 5, no. 11, pp. 831-843, 2004.

[6] J. Y. Song, K. Demuth, and J. Morgan, "Effects of the acoustic properties of infant-directed speech on infant word recognition," The Journal of the Acoustical Society of America, vol. 128, no. 1, pp. 389-400, 2010.

[7] H. M. Liu, P. K. Kuhl, and F. M. Tsao, "An association between mothers' speech clarity and infants' speech discrimination skills," Developmental science, vol. 6, no. 3, pp. F1-F10, 2003.

[8] K. T. Englund and D. M. Behne, "Infant directed speech in natural interaction—Norwegian vowel quantity and quality," Journal of psycholinguistic research, vol. 34, no. 3, pp. 259-280, 2005.

[9] T. Benders, "Nature's distributional-learning experiment," Citeseer, 2013.

[10] N. X. Rattanasone, D. Burnham, and R. G. Reilly, "Tone and vowel enhancement in Cantonese infant-directed speech at 3, 6, 9, and 12 months of age," Journal of Phonetics, vol. 41, no. 5, pp. 332-343, 2013.

[11] Burnham, D., Best, C. (Cutler, A., Johnson, E., Kalashnikova, M.) Origins of Phonology and Lexicon: Abstract representations before 6 months Australian Research Council Discovery Grant, DP00025567.

[12] A. Cristia and A. Seidl, "The hyperarticulation hypothesis of infant-directed speech," Journal of child language, vol. 41, no. 4, pp. 913-934, 2014.

[13] B. McMurray, K. A. Kovack-Lesh, D. Goodwin, and W. McEchron, "Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence?," Cognition, vol. 129, no. 2, pp. 362-378, 2013.

[14] M. Kalashnikova, C. Carignan, and D. Burnham, "The origins of babytalk: smiling, teaching or social convergence?," Royal Society open science, vol. 4, no. 8, p. 170306, 2017.

[15] M. Kalashnikova, U. Goswami, and D. Burnham, "Mothers speak differently to infants at-risk for dyslexia," Developmental science, vol. 21, no. 1, p. e12487, 2018.

[16] C. Kitamura, R. Panneton, and C. T. Best, "The development of language constancy: Attention to native versus nonnative accents," Child Development, vol. 84, no. 5, pp. 1686-1700, 2013.

[17] P. Roach, "British English: received pronunciation," Journal of the International Phonetic Association, vol. 34, no. 2, pp. 239-245, 2004.

[18] F. Cox and J. Fletcher, Australian English pronunciation and transcription. Cambridge University Press, 2017.

[19] P. Boersma, "Praat: doing phonetics by computer (Version 5.1. 05)," http://www. praat. org/, 2009.

[20] Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. Journal of Phonetics, 34(4), 458–484.

[21] R Core Team, R: The R project for statistical computing. [Online]. Available: https://www.r-project.org, 2022.

[22] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," J. Stat. Soft., vol. 67, no. 1, pp. 1–48, Oct. 2015.

[23] H. Wickham, "Programming with ggplot2," in ggplot2: Springer, 2016, pp. 241-253.

[24] C. Lam and C. Kitamura, "Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation," Developmental Science, vol. 15, no. 2, pp. 212-221, 2012.

[1] Note that /r/ is not rhoticized in Australian English. The word *chook* is a common informal name for 'chicken' in Australian English.