

Research Article

Phonetic Modification of Vowel Space in Storybook Speech to Infants up to 2 Years of Age

Evamarie B. Burnham,^a Elizabeth A. Wieland,^a Maria V. Kondaurova,^b
J. Devin McAuley,^a Tonya R. Bergeson,^b and Laura C. Dilley^a

Purpose: A large body of literature has indicated vowel space area expansion in infant-directed (ID) speech compared with adult-directed (AD) speech, which may promote language acquisition. The current study tested whether this expansion occurs in storybook speech read to infants at various points during their first 2 years of life.

Method: In 2 studies, mothers read a storybook containing target vowels in ID and AD speech conditions. Study 1 was longitudinal, with 11 mothers recorded when their infants were 3, 6, and 9 months old. Study 2 was cross-sectional, with 48 mothers recorded when their infants were 3, 9, 13,

or 20 months old ($n = 12$ per group). The 1st and 2nd formants of vowels /i/, /a/, and /u/ were measured, and vowel space area and dispersion were calculated.

Results: Across both studies, 1st and/or 2nd formant frequencies shifted systematically for /i/ and /u/ vowels in ID compared with AD speech. No difference in vowel space area or dispersion was found.

Conclusions: The results suggest that a variety of communication and situational factors may affect phonetic modifications in ID speech, but that vowel space characteristics in speech to infants stay consistent across the first 2 years of life.

Over the past two decades, a number of studies have investigated the acoustic–phonetic characteristics of infant-directed (ID) speech compared with adult-directed (AD) speech. A central finding in these studies is that the area of the vowel space triangle, formed by the first (F1) and second (F2) formants of point vowels /i/, /a/, and /u/, is larger in ID than AD speech (Burnham, Kitamura, & Vollmer-Conna, 2002; Cristia & Seidl, 2013; Kondaurova, Bergeson, & Dilley, 2012; Kuhl et al., 1997; Liu, Kuhl, & Tsao, 2003; Liu, Tsao, & Kuhl, 2009; Uther, Knoll, & Burnham, 2007; Xu, Burnham, Kitamura, & Vollmer-Conna, 2013). This larger vowel space triangle area has been suggested to reflect hyperarticulation of the point vowels /i/, /a/, and /u/ to articulatory positions that are more phonologically contrastive and clearer in ID as compared with AD speech (Kuhl et al., 1997). Such modifications are thought to facilitate phonological category

learning and promote language acquisition in infants (Burnham et al., 2002; Cristia & Seidl, 2013; Kondaurova et al., 2012; Kuhl et al., 1997; Liu et al., 2003; Liu et al., 2009; Uther et al., 2007).

Accumulated evidence from recent studies has cast doubts on the characterization of ID speech as “hyper-articulation” per se (Benders, 2013; Cristia & Seidl, 2013; Englund & Behne, 2006; Kondaurova et al., 2012; Lam & Kitamura, 2010, 2012). Several studies have demonstrated a lack of expansion of the acoustic vowel space area and/or no enhancement of the distance between vowel categories in ID as compared with AD speech (Benders, 2013; Cristia & Seidl, 2013; Englund & Behne, 2006; Kondaurova et al., 2012; Xu Rattanasone, Burnham, & Reilly, 2013). Other studies have shown that the degree of change in the vowel space area in ID compared with AD speech depends on the infant’s hearing status, both simulated and naturally occurring. For example, Lam and Kitamura (2012) used a closed-circuit television setup to test mothers’ interactions with children with normal hearing and found that the amount of vowel space expansion in mothers’ speech was dependent on whether and how well the infant could hear the mother. It was suggested that this modulation of vowel space area was due to a decrease in infant feedback when the infant could not hear his or her mother. A case study of

^aMichigan State University, East Lansing

^bIndiana University School of Medicine, Indianapolis

Correspondence to Laura C. Dilley: ldilley@msu.edu

Editor: Jody Kreiman

Associate Editor: Megha Sundara

Received August 1, 2013

Revision received May 2, 2014

Accepted December 27, 2014

DOI: 10.1044/2015_JSLHR-S-13-0205

Disclosure: The authors have declared that no competing interests existed at the time of publication.

a mother interacting with her twins, one with a hearing impairment and one with normal hearing, demonstrated vowel space area expansion in speech to the child with normal hearing but not to the child with hearing impairment (Lam & Kitamura, 2010). Similarly, data from Norwegian and Dutch speech research have shown a lack of vowel space expansion in ID as compared with AD speech (Benders, 2013; Englund & Behne, 2006). This evidence suggests that vowel space area expansion is not an inherent characteristic of all ID speech, but instead, features that differentiate ID from AD speech depend on a variety of communicative and situational factors.

Consequently, it is not clear under what conditions caregivers might or might not produce an expanded vowel space. Because vowel space expansion is thought to promote language acquisition, it is of interest to examine conditions in which ID speech has clear didactic functions. One such condition is storybook speech, an activity valued by educators and parents and recommended in early-childhood policy documents as a means of encouraging language development and literacy (Bowman, Donovan, & Burns, 2000; Evans & Saint-Aubin, 2005; Fitzgerald, Spiegel, & Cunningham, 1991; Snow, Burns, & Griffin, 1998). The use of scripted storybook speech affords a strategic advantage in this study design, namely the ability to ensure a large number of point vowel tokens (/i/, /a/, and /u/) for examination as a function of speech style, as well as control of syntactic and semantic variation.

A second important motivation for the present study was to investigate how the ID vowel space might change as a function of infants' age (Bernstein Ratner, 1984; Kitamura & Burnham, 2003; Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002; Stern, Spieker, Barnett, & MacKain, 1983). In a pioneering study of vowel articulation in speech to infants, Bernstein Ratner (1984) found patterns suggestive of both enhanced overall vowel clarity in ID speech as well as a shift in ID vowel production depending on the infant's stage of development. Similarly, Xu Rattanasone et al. (2013) found evidence of a possible increase in vowel space area as a function of infant age in speech to children between 3 and 12 months; although this difference was nonsignificant, they suggested that a more definite pattern might emerge when including speech to children who were over a year old. A possible change in ID vowel articulation based on infant age is also supported by research showing shifts in infant preferences for particular features of ID speech at different ages (Hayashi, Tamekawa, & Kiritani, 2001; Kuhl et al., 2008). These shifts in developmental preference may affect the maternal input in a way that will cause the nature of modification to the ID vowel space to change over time (Smith & Trainor, 2008).

Many studies have clarified the nature of vowel modification in ID speech, but most have examined one or maximally two isolated time points (Benders, 2013; Burnham et al., 2002; Cristia & Seidl, 2013; Kondaurova et al., 2012; Kuhl et al., 1997; Lam & Kitamura, 2012; Liu et al., 2003; Liu et al., 2009; Xu et al., 2013). Other relevant studies have either sampled over a relatively short time period in

infant development (Englund & Behne, 2005, 2006) or else have lumped a sizeable infant age range into a single time point (Uther et al., 2007).

The goal of the present study was to test the generalizability of vowel hyperarticulation in ID speech across speech contexts. This study focused on speech arising from the context of storybook reading. Specifically, the present studies compared phonetic characteristics of the vowel space area between ID and AD conditions in storybook speech. If caregivers produce ID speech to help instruct their children about the phonological categories of the language, then enhanced segmental characteristics (i.e., vowel space expansion) should be evident in the phonetic characteristics of storybook speech read to infants. However, if vowel hyperarticulation is modulated by infant feedback, as Lam and Kitamura (2012) suggested, vowel space expansion may be less likely to occur in a storybook condition, in which mothers may be focusing their attention on the book rather than on their infants. Therefore, the current study tested whether ID speech shows hyperarticulation in a storybook reading context.

In addition, we considered how the ID vowel space might change over a relatively wide span of early development, from very early (approximately 3 months of age) to relatively late (approximately 20 months of age). In order to obtain representative samples of mothers' speech at appropriate intervals, we included both a longitudinal study investigating speech at 3, 6, and 9 months, and a cross-sectional study investigating behavior at 3, 9, 13, or 20 months. A significant strength of the present research involves the use of two distinct studies across age groups employing complementary research designs. In particular, Study 1 used a longitudinal design to minimize extraneous variability that may arise from distinct participant groups across age intervals, whereas Study 2 used a cross-sectional design to avoid potentially confounding familiarity effects associated with within-subjects longitudinal designs. The use of a standardized text for read storybook speech facilitated examination of speech characteristics across multiple age groups while ensuring experimental control over context and speech materials.

Study 1

The aim of Study 1 is to compare acoustic-phonetic characteristics (F1 and F2 frequencies) of three point vowels /i/, /a/, and /u/ in ID and AD speech produced by mothers in an ID condition to the same infants at 3, 6, and 9 months of age and in an AD condition at corresponding time intervals.

Method

Participants

Eleven mother-infant dyads were recruited from the local community in Indianapolis, Indiana. Each mother and her infant participated in three separate sessions when the infant was approximately 3, 6, and 9 months of age. The

mean age of the infants (four girls, seven boys) at the first interval was 3.0 months ($SD = 0.3$; range = 2.3–3.4), at the second interval was 5.9 months ($SD = 0.4$; range = 5.0–6.5), and at the third interval was 9.0 months ($SD = 0.2$; range = 8.8–9.5). All mothers were native speakers of American English with self-reported normal hearing who grew up in the Midwestern United States and were paid \$10 per visit. This research and the recruitment of human subjects were approved by the Indiana University Institutional Review Board.

Procedure

Recordings

Mothers were digitally recorded reading in a double-walled, copper-shielded sound booth (Industrial Acoustics Company, Bronx, NY). The speech was recorded in one of two ways: The initial system used an Audio-Technica ES933/H hypercardioid microphone (Audio-Technica, Leeds, UK) powered by a phantom power source and linked to an amplifier (DSC-240; Daqscribe, Centennial, CO) and a Sony DTC-690 digital audio tape recorder (Sony, Tokyo, Japan). The equipment was updated partway through this longitudinal project to an SLX Wireless Microphone System (Shure, Niles, IL). This system included an SLX1 Bodypack transmitter with a built-in microphone and a wireless receiver SLX4, which was connected to a Canon 3CCD Digital Video Camcorder GL2, NTSC (Canon, Melville, NY) and recorded the speech samples directly onto a Mac computer (OSX Version 10.4.10; Apple, Inc., Cupertino, CA) via Hack TV (Version 1.11) software. No systematic differences were found across recording sessions or participant groups in terms of recording technology. Recordings were made at a sampling rate of 22050 Hz with 16-bit quantization rate.

Token Identification

Mothers read from a storybook specifically constructed to contain key words with the target /i/, /a/, and /u/ vowels (see Appendix A). In the ID speech condition, mothers were asked to read to their infants at each of the three sessions; recordings ranged from 2.0 to 4.0 minutes ($M = 2.7$, $SD = 0.5$). The infants were present in the room with their mothers for the duration of the ID session. In the AD condition, mothers were asked to read aloud as if to another adult at each of the three sessions; recordings ranged from 1.3 to 2.2 minutes ($M = 1.7$, $SD = 0.3$). The order of ID and AD recordings was counterbalanced across mothers.

In total, 66 recordings were collected across both conditions (ID = 11 mothers \times 3 sessions, AD = 11 mothers \times 3 sessions). A target of 20 tokens of each vowel from the words in the storybook was included in the analysis; if more than 20 were produced, the tokens were randomly selected. On average, 15.1 ($SD = 2.1$) tokens of each vowel (/i/, /a/, and /u/) in the ID speech condition and 15.8 ($SD = 1.7$) tokens of each vowel (/i/, /a/, and /u/) in the AD speech condition were analyzed for each speaker at each interval. A total of 1,499 tokens in the ID speech condition and 1,559 tokens in the AD speech condition were analyzed. The Praat 5.0.21 editor (Boersma & Weenink, 2012) and

MATLAB (MathWorks, 2009) software were used to identify and segment out each vowel in recorded speech based on a combination of waveform and spectral cues. Due to a vowel merger between /a/ and /ɔ/ in progress in Indiana (Labov, Ash, & Boberg, 2006), it was impossible to reliably determine for a given speaker whether a particular token of a low back vowel reflected a single phonemic category or two, given the expectation of substantial within-category variation in the vowel space for the ID register overall. As a result, all instances of low back vowels were treated as a single category.

Acoustic Analysis

Formant frequencies. Phonetic analysts trained in formant analysis first identified the onset and offset of each randomly selected vowel token via visual inspection of spectrogram and waveform information using segmentation criteria established for the Buckeye Corpus (Pitt et al., 2007). Measurements of the first (F1) and second (F2) formants were then taken at the vowel midpoint using a combination of spectral slices, visual inspection of spectrograms, and linear predictive coding (LPC) estimates derived from Praat (Boersma & Weenink, 2012) and FormantMeasurer software in MATLAB (Morrison & Nearey, 2011; MathWorks, 2009); all F1 and F2 measurements were checked by hand for correctness. Analysts identified individual tokens of the target vowels as usable if the first two formants were reasonably clear and measurements fell within an expected range of the mean, plus or minus three standard deviations, as determined by mean formant values for female talkers across multiple studies tabulated in Kent and Read (1992). Tokens that fell outside the expected range had strongly stratified harmonics, or for which F1 or F2 could not be determined (i.e., due to high F0, coarticulation, poor sound quality, or background noise), were checked by one of the authors for usability before being included in or excluded from the analysis. Tokens with high F0 (i.e., greater than 350 Hz) were often excluded due to the quantization of the spectrum and thus greater variability and unreliability of formant measurements in high-F0 tokens (Vallabha & Tuller, 2002). If a randomly selected token of a given vowel was excluded for any reason, it was replaced by another randomly selected token of that vowel from among the remaining tokens produced by that mother in the same speech condition. A total of 18.5% of selected tokens were excluded for various reasons (5.3% AD condition, 28.5% ID condition); of these, 4.6% had F0 over 350 Hz (0.0% AD condition, 5.3% ID condition).

Formant values in Hz were converted to mel units. The mel scale is based on psychophysical studies of pitch distance and reflects human perception of frequency more directly than linear Hz. The relationship between the mel scale and Hz is a nonlinear, strictly monotonic increasing function, such that above 500 Hz, larger and larger intervals are judged by listeners to produce equal pitch increments. The mel scale has been used in many prior studies of vowel space and formant characteristics (e.g., Bradlow, Torretta, & Pisoni, 1996; Englund & Behne, 2005; Kuhl

et al., 1997; Lam & Kitamura, 2010, 2012; Xu et al., 2013; Xu Rattanasone et al., 2013). The following equation was used for the Hz to mels conversion (Fant, 1973; in Bradlow et al., 1996):

$$mels = \frac{1000}{\log 2} \times \log \left(\frac{Hz}{1000} + 1 \right) \quad (1)$$

The mels conversion provided the basis for all analyses reported. The means and standard deviations of F1 and F2 were determined for each speaker in both ID and AD speech conditions.

Vowel space area. Vowel space triangles were constructed in an x - y plane, where the average F1 and F2 values of /i/, /a/, and /u/ vowels were the respective x and y coordinates of the corners. The area of the resultant triangles in both ID and AD conditions was calculated using the following equation (Baker, 1885; Liu et al., 2003; Weisstein, 2014):

$$Area = \frac{|(F1_i \times (F2_a - F2_u) + F1_a \times (F2_u - F2_i) + F1_u \times (F2_i - F2_a))|}{2} \quad (2)$$

Vowel space dispersion. Previous research on AD speech has identified vowel space dispersion as a good index of speech clarity (Bradlow et al., 1996). The vowel space dispersion is calculated by measuring the distance of each token from a central point in the talker's vowel space. This measure provides an indication of the overall expansion/compaction of the set of vowel tokens from each participant, and detects fine-grained individual differences in acoustic-phonetic characteristics (Bradlow et al., 1996). By capturing a different aspect of vowel production characteristics than the traditional Heron method (Kuhl et al., 1997; Neel, 2008), this metric helps to provide an assessment of vowel clarity. Vowel space dispersion was calculated by finding the centroid of each speaker's vowel space triangle and averaging the distances of the individual tokens from the centroid (see Appendix B for further detail; Bradlow et al., 1996).

Reliability

The formants of any vowel token whose F1 or F2 was two or more standard deviations away from a given participant's mean F1 or F2, respectively, were checked by hand to ensure accuracy. In addition, trained analysts re-measured a random selection of 5% of the tokens used in each speech sample for an analysis of interrater reliability. The percentage difference (Δ_i) between the first rater's measurement (r_1) and the second rater's measurement (r_2) was calculated using the following equation (Kuhl et al., 1997):

$$\Delta_i = (|r_1 - r_2| / r_1) \times 100\% \quad (3)$$

The average interrater percentage difference was 5.7% ($SD = 4.8$). This is in line with reliability reported in previous studies and indicates high interrater reliability (e.g., Kuhl et al., 1997).

Results and Discussion

Formant Frequencies

The means and standard deviations for F1 and F2 frequencies of each point vowel are shown in Table 1. A 2 (Speech Style: ID vs. AD speech) \times 3 (Interval: 3, 6, and 9 months) repeated-measures analysis of variance (ANOVA) was conducted with speech style and interval as within-subjects factors for each formant (F1, F2) of each point vowel (/i/, /a/, and /u/).

The results for /i/ demonstrated a significant effect of speech style with F2 higher in ID compared with AD speech, $F(1, 10) = 34.424, p < .001, \eta_p^2 = 0.775$. A higher F2 corresponds to a generally more advanced or fronted tongue position according to standard acoustic-phonetic modeling assumptions and empirical findings (Stevens, 2000); thus, our data suggest a more advanced/fronted tongue position in ID ($M = 1898$ mels, $SD = 58$ mels) than AD ($M = 1860$ mels, $SD = 49$ mels) speech. No main effect of interval or interaction between interval and speech style was found for F2. The results also demonstrated a significant effect of speech style with F1 higher in ID compared with AD speech, $F(1, 10) = 9.556, p = .011, \eta_p^2 = 0.489$. Since a higher F1 is generally assumed to correspond to a lower tongue position (Stevens, 2000), our data suggest there is a lower tongue position in ID ($M = 539$ mels, $SD = 35$ mels) than AD ($M = 496$ mels, $SD = 49$ mels) speech. No main effect of interval or interaction between interval and speech style were found for F1.

The results for /a/ demonstrated a significant effect of interval on F2, $F(2, 20) = 3.699, p = .043, \eta_p^2 = 0.27$. However, no pairwise comparisons of interval were significant in post hoc t tests after a Bonferroni correction to $p = (.05/3) = .017$ (3 months: $M = 1186$ mels, $SD = 55$ mels; 6 months: $M = 1176$ mels, $SD = 63$ mels; 9 months: $M = 1202$ mels, $SD = 48$ mels); this lack of post hoc significance likely reflects the conservative nature of Bonferroni tests (i.e., their relatively high Type II error rate). No main effect of speech style or interaction between interval and speech style were found for F2. No significant main effects or interaction were found for F1.

The results for /u/ demonstrated a significant effect of speech style, with F2 lower in ID compared with AD speech, $F(1, 10) = 5.552, p = .040, \eta_p^2 = 0.357$; this may indicate a more retracted/backed tongue position in ID ($M = 1169$ mels, $SD = 91$ mels) than AD ($M = 1190$ mels, $SD = 83$ mels) speech. No main effect of interval or interaction between interval and speech style were found for F2. The results also demonstrated a significant effect of speech style, with F1 higher in ID compared with AD speech, $F(1, 10) = 27.752, p < .001, \eta_p^2 = 0.735$, which may indicate a lower tongue position in ID ($M = 543$ mels, $SD = 30$ mels) than AD ($M = 518$ mels, $SD = 28$ mels) speech.

Vowel Space Area and Vowel Space Dispersion

We first examined whether an overall difference in vowel space area or vowel space dispersion existed. The means and standard deviations of the vowel space area and

Table 1. Mean (SD) F1 and F2 frequencies in mels for the point vowels /i/, /a/, and /u/ in infant-directed (ID) and adult-directed (AD) speech at three longitudinal intervals (Study 1) and at four cross-sectional intervals (Study 2).

Interval	Speech Style	/i/		/a/		/u/	
		F1	F2	F1	F2	F1	F2
Study 1							
3 months	AD	483 (41)	1861 (41)	879 (41)	1185 (51)	509 (25)	1194 (74)
	ID	544 (36)	1910 (57)	877 (60)	1187 (61)	541 (24)	1170 (60)
6 months	AD	501 (48)	1866 (46)	845 (44)	1182 (53)	518 (29)	1189 (77)
	ID	540 (39)	1888 (48)	861 (44)	1170 (73)	546 (24)	1163 (94)
9 months	AD	503 (59)	1852 (61)	841 (51)	1198 (40)	528 (29)	1188 (101)
	ID	533 (33)	1897 (70)	867 (69)	1206 (56)	541 (42)	1174 (118)
Study 2							
3 months	AD	534 (44)	1878 (60)	849 (36)	1158 (41)	527 (39)	1195 (72)
	ID	551 (31)	1895 (72)	843 (56)	1165 (78)	541 (29)	1175 (63)
9 months	AD	500 (65)	1884 (65)	851 (60)	1186 (62)	523 (26)	1173 (104)
	ID	506 (68)	1900 (61)	857 (54)	1176 (83)	543 (25)	1204 (135)
13 months	AD	504 (52)	1872 (56)	843 (52)	1168 (74)	515 (32)	1185 (33)
	ID	526 (32)	1915 (48)	837 (55)	1183 (58)	528 (39)	1179 (52)
20 months	AD	518 (53)	1850 (73)	826 (38)	1161 (44)	516 (37)	1182 (54)
	ID	546 (42)	1885 (57)	841 (54)	1167 (64)	547 (30)	1185 (57)

dispersion values for ID and AD speech are reported in Table 2, and Figure 1 shows vowel space triangles. A 2 (Speech Style: ID vs. AD speech) × 3 (Interval: 3, 6, and 9 months) repeated measures ANOVA was conducted separately for vowel space area and vowel space dispersion with speech style and interval as within-subjects factors.

The difference in mean vowel space area for ID ($M = 118697 \text{ mels}^2$, $SD = 26351 \text{ mels}^2$) compared with AD ($M = 113166 \text{ mels}^2$, $SD = 26405 \text{ mels}^2$) speech was not significant as a function of interval or speech style, and there was no significant interaction between these variables. Also, the difference in mean vowel space dispersion for ID ($M = 395 \text{ mels}$, $SD = 32 \text{ mels}$) and AD ($M = 384 \text{ mels}$, $SD = 28 \text{ mels}$) speech was not significant as a function of interval or speech style, and there was no significant interaction between these variables.¹

Summary

Overall, Study 1 showed a reliable shift in F1 and F2 for both /i/ and /u/ in ID as compared with AD speech, but no effect of speech style on /a/ formant values; there was also no significant expansion of vowel space area or dispersion in ID as compared with AD speech.² These results provide little support for across-the-board enhancements to phonological contrast in ID speech compared with AD speech. The lack of a difference in vowel space area and dispersion between speech styles for this study may be due to the use

¹Percent change of vowel space area and dispersion were also calculated using the formula $\% \Delta_X = \left[\frac{X_{ID} - X_{AD}}{X_{AD}} \right] \times 100\%$, where X represents the percentage for the metric in question (i.e., either vowel space area or dispersion) and the subscript indicates the Speech Style. Results of this alternative analysis were nearly identical to the analysis of raw vowel space area and raw vowel space dispersion.

²The data were also analyzed using raw Hz as a metric; statistical results using ANOVA were very similar.

of a storybook for elicitation, which has not been frequently employed in previous ID vowel space studies.

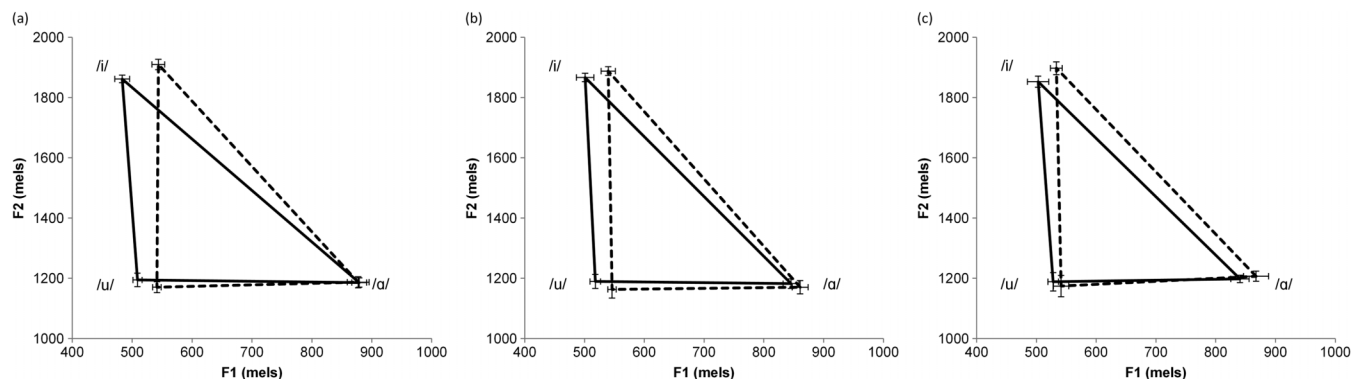
Study 2

The aim of Study 2 was to extend the findings from Study 1 comparing acoustic-phonetic characteristics (F1 and F2 frequencies) in ID and AD speech to a larger number of participants ($n = 48$) over a more extended period of time from young infant (3 months) to the toddler stage (20 months). In addition, conducting a separate cross-sectional research study permitted the opportunity to obtain converging evidence for vowel space modifications in storybook speech using a research design complementary to Study 1.

Table 2. Mean (SD) vowel space area in mels² and vowel space dispersion in mels for infant-directed (ID) and adult-directed (AD) speech at three longitudinal intervals (Study 1) and at four cross-sectional intervals (Study 2).

Study	Interval	AD	ID
Study 1			
Area	3 months	123399 (22508)	123574 (25280)
	6 months	111763 (24742)	113739 (20865)
	9 months	104335 (30153)	118778 (33082)
Dispersion	3 months	393 (27)	395 (24)
	6 months	386 (16)	399 (35)
	9 months	374 (36)	390 (37)
Study 2			
Area	3 months	110990 (23513)	109293 (25576)
	9 months	117860 (26098)	109247 (22875)
	13 months	113386 (19883)	114137 (26531)
	20 months	104631 (22937)	103975 (30412)
Dispersion	3 months	389 (35)	388 (33)
	9 months	382 (29)	389 (28)
	13 months	391 (29)	403 (23)
	20 months	379 (42)	380 (47)

Figure 1. Infant-directed (dashed line) and adult-directed (solid line) vowel space triangles based on mean F1 and F2 values for speech to participants at the (a) 3-month, (b) 6-month, and (c) 9-month longitudinal intervals in Study 1. Error bars show mean \pm 1 SEM.



Method

Participants

Forty-eight mother–infant dyads were recruited for participation from the same subject pool as Study 1. Each mother and her infant participated in a single session when the infant was approximately 3, 9, 13, or 20 months of age. The mean age of the infants at the first interval was 3.1 months ($SD = 0.4$; range = 2.5–4.1; four girls, eight boys), at the second interval was 9.0 months ($SD = 0.4$; range = 8.3–9.9; four girls, eight boys), at the third interval was 12.8 months ($SD = 0.5$; range = 12.1–13.8; four girls, eight boys), and at the fourth interval was 20.4 months ($SD = 0.9$; range = 18.7–21.8; six girls, six boys). Recordings from six participants in Study 1 were included in Study 2 from a single temporal interval (i.e., exactly one ID and one AD recording per mother), thereby balancing sample sizes across groups and increasing statistical power. Each of the 48 mothers was assigned to exactly one group based on the age of her child at the time of the selected recording. This meant that, although each mother was included in the analysis at only one time point, that time point was not always the first time the mother read the storybook to her infant. Research approval and participant reimbursement were identical to Study 1.

Procedure

Recordings

The recording procedure was identical to Study 1. In total, there were 96 recordings across ID and AD speech conditions. ID sessions ranged from 1.6 to 3.9 minutes ($M = 2.6$, $SD = 0.5$). AD sessions ranged from 1.3–2.2 minutes ($M = 1.7$, $SD = 0.2$). The order of ID and AD recordings was counterbalanced across mothers.

Token Identification

The token identification procedure and inclusion criteria were identical to Study 1. On average, 14.3 ($SD = 3.2$) tokens of each vowel (/i/, /a/, and /u/) in the ID speech condition and 15.2 ($SD = 2.6$) tokens of each vowel (/i/, /a/, and /u/) in the AD speech condition were analyzed for each

speaker. A total of 2,062 tokens in the ID speech condition and 2,194 tokens in the AD speech condition were analyzed.

Acoustic Analysis

Analysis procedures for measuring vowel space area, vowel space dispersion, and F1 and F2 frequencies were identical to Study 1. A total of 11.2% of selected tokens were excluded for various reasons (4.7% AD condition, 17.1% ID condition); of these, 40.9% were excluded for high F0 (23.9% AD condition, 45.3% ID condition).

Reliability

Interrater reliability analysis procedures were identical to Study 1. The average interrater percentage difference was 8.9% ($SD = 8.5$), indicating high reliability consistent with that of prior studies (e.g., Kuhl et al., 1997).

Results and Discussion

Formant Frequencies

The means and standard deviations for F1 and F2 frequencies of each point vowel are shown in Table 1. A 2 (Speech Style: AD, ID) \times 4 (Interval: 3, 9, 13, or 20 months) mixed measures ANOVA was conducted with speech style as a within-subjects factor and interval as a between-subjects factor for each formant (F1, F2) of each point vowel (/i/, /a/, and /u/).³ The results for /i/ demonstrated a significant effect of speech style, with F2 higher in ID compared with AD speech, $F(1, 44) = 19.042$, $p < .001$, $\eta_p^2 = 0.302$, which may indicate a more advanced/fronted tongue position in ID ($M = 1899$ mels, $SD = 59$ mels) than AD ($M = 1871$ mels, $SD = 63$ mels) speech. No main effect of interval or interaction between interval and speech style was found for F2. The results also demonstrated a significant effect of speech style, with F1 higher in ID compared with AD speech, $F(1, 44) = 6.956$, $p = .012$, $\eta_p^2 = 0.137$, which may indicate

³ANCOVA analyses with number of prior recording sessions as a covariate resulted in similar results to the ANOVA looking at the production of formant frequencies in the point vowels /i/, /a/, and /u/.

a lower tongue position in ID ($M = 532$ mels, $SD = 48$ mels) than AD ($M = 514$ mels, $SD = 54$ mels) speech. No main effect of interval or interaction between interval and speech style were found for F1. The results for /a/ demonstrated no main effects of speech style, interval, or interaction between interval and speech style for F1 or F2. The results for /u/ demonstrated a significant effect of speech style with F1 higher in ID compared with AD speech, $F(1, 44) = 20.092$, $p < .001$, $\eta_p^2 = 0.313$, which may indicate a lower tongue position in ID ($M = 540$ mels, $SD = 31$ mels) than AD ($M = 520$ mels, $SD = 33$ mels) speech. No main effect of interval or interaction between interval and speech style was found for F1. No significant main effect or interaction was found for F2.

Vowel Space Area and Dispersion

We first examined whether an overall difference in vowel space area or vowel space dispersion existed. The means and standard deviations of the vowel space area and dispersion values for ID and AD speech are reported in Table 2, and Figure 2 shows vowel space triangles. A 2 (Speech Style: AD, ID) \times 4 (Interval: 3, 9, 13, or 20 months) mixed measures ANOVA was conducted separately for vowel space area and vowel space dispersion with speech style as a within-subjects factor and interval as a between-subjects factor.

The difference in mean vowel space area for ID ($M = 109163$ mels², $SD = 25883$ mels²) compared with AD ($M = 111717$ mels², $SD = 22973$ mels²) speech was not significant as a function of interval or speech style, and there was no significant interaction between these variables. The difference in mean vowel space dispersion for ID ($M = 390$ mels, $SD = 34$ mels) compared with AD ($M = 385$ mels, $SD = 34$ mels) speech was also not significant as a function of interval or speech style, and there was no significant interaction between these variables.⁴

Summary

Overall, the results of Study 2 were very similar to those of Study 1. A reliable shift in F1 and F2 for /i/ and in F1 for /u/ was found in ID as compared with AD speech, but there was no effect of speech style on /a/ formant values; additionally, no significant expansion of vowel space area or vowel space dispersion was found in ID as compared with AD speech.⁵ The lack of a difference in

our study for vowel space area and vowel space dispersion between speech styles may again be due to the use of a storybook for elicitation, which has not been employed frequently in previous ID vowel space studies. The pattern of results was very similar over time, consistent with Study 1.

General Discussion

The current study tested the generalizability of vowel hyperarticulation in ID speech by examining vowel space characteristics of storybook speech directed to infants and to adults. In addition, the current study examined vowel space characteristics of ID speech over a wide range of infant ages in order to gain a clearer picture of how ID speech changes over time. Measurements were conducted using scripted storybook speech in ID and AD conditions over approximately the first 9 months (Study 1) or the first 20 months (Study 2) of infant life.

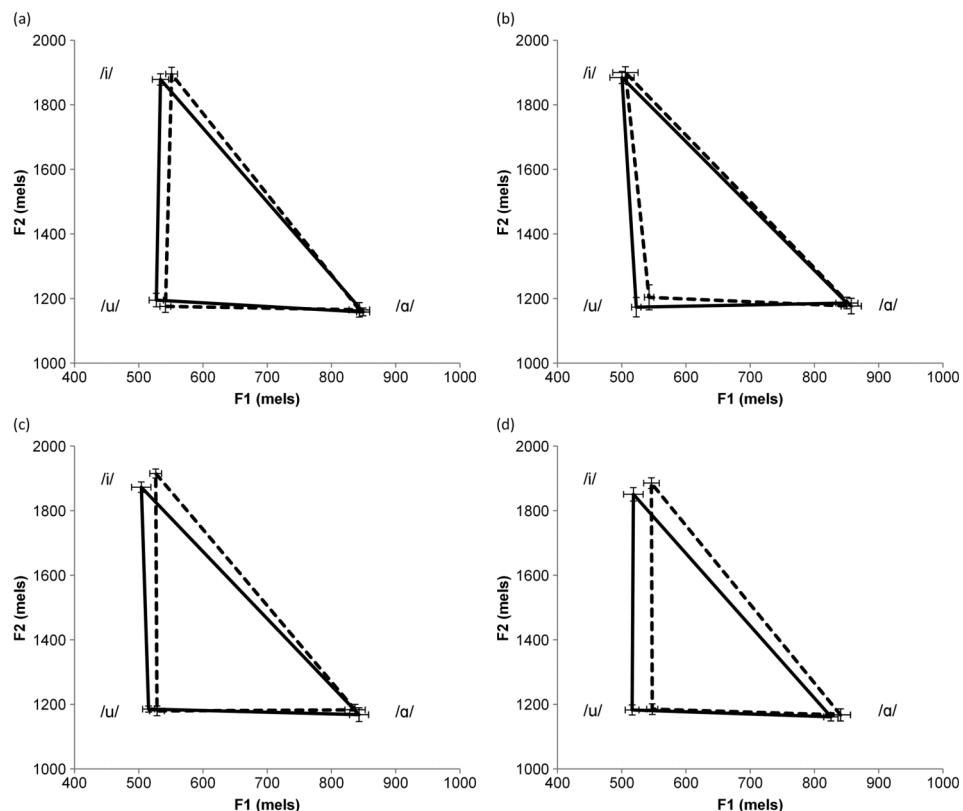
The results of both Study 1 and Study 2 indicated systematic shifts of F1 and/or F2 frequencies for /i/ and /u/ vowels in ID as compared with AD speech. However, vowel triangle areas overall were not larger in ID compared with AD speech. In addition, vowel clarity, as indexed by vowel space dispersion (Bradlow et al., 1996), was not enhanced in ID compared with AD speech.

These results suggest limitations to the generalizability of hyperarticulation in ID speech to a storybook context. Both of our studies demonstrated higher F1 frequencies for /i/ and /u/ vowels in ID as compared with AD speech. In contrast, across-the-board hyperarticulation would be expected to result in a *lower* F1 for these vowels in order to maximize the distance between vowels in ID speech (Kuhl et al., 1997). The increase in F2 formant values for the /i/ vowel and the decrease in F2 formant values for the /u/ vowel are consistent with findings from previous research suggesting a more advanced/fronted tongue position for /i/ and more retracted/backed tongue position for /u/ in ID as compared with AD speech (Bernstein Ratner, 1984; Kondaurova et al., 2012; Kuhl et al., 1997). For the /a/ vowel, no effect of speech style or interval was found for either F1 or F2 frequencies, despite evidence from previous studies that predicted an increase in their values (Bernstein Ratner, 1984; Kuhl et al., 1997). The results demonstrate limited evidence for hyperarticulation of the point vowels /i/ and /u/, specifically with regard to the F2 dimension. However, when these results are considered alongside the raised F1 for /i/ and /u/, as well as the lack of change in /a/ formant values, the overall picture does not correspond to across-the-board hyperarticulation, which would be expected to produce a lower F1 for the high vowels /i/ and /u/, as well as a possibly higher F1 for the low vowel /a/. However, these results do suggest that the distribution of vowels in acoustic vowel space depends on the speech style, even for a speech context where there is no overall expansion of the vowel space area (Benders, 2013; Bernstein Ratner, 1984; Englund & Behne, 2005; Kondaurova et al., 2012; Kuhl et al., 1997).

⁴Percent change of vowel space area and vowel space dispersion were also calculated following the same method as for Study 1. Results of this alternative analysis were identical to the analysis of raw vowel space area and raw vowel space dispersion.

⁵The data were also analyzed using raw Hz, and the ANOVA results were nearly identical. The only difference was seen in the vowel space dispersion, which went from being nonsignificant ($p = .056$) to ID speech being significantly greater than AD speech ($p = .006$). This is likely due to the fact that mels is a logarithmic scale representing human perception, so that a significant difference in Hz in the highest frequencies may not translate to a significantly different perception by the human ear.

Figure 2. Infant-directed (dashed line) and adult-directed (solid line) vowel space triangles based on mean F1 and F2 values for speech to participants at the (a) 3-month, (b) 9-month, (c) 13-month, and (d) 20-month cross-sectional intervals in Study 2. Error bars show mean \pm 1 SEM.



Englund and Behne (2005) attributed findings of formant raising in ID speech to increased smiling behavior in interaction with infants. This interpretation is consistent with data showing that infants respond more to positive affect, and that mothers have higher positive affect in ID than AD speech (Kitamura & Burnham, 1998, 2003). Similarly, Benders (2013) found evidence supporting a correlation between F2 raising and higher positive affect in Dutch ID speech. Therefore, it seems possible that the increase in formant frequencies in the current study could also be accounted for by mothers' smiling. Although an increase in both F1 and F2 was found for only two out of the three vowels examined here, previous studies likewise have found an overall increase in vowel formant values in smiled speech, but with effects varying for specific vowel-formant combinations (Fagel, 2009; Kienast & Sendmeier, 2000; Tartter, 1980; Tartter & Braun, 1994). This variability in results may be an indication of large interindividual and/or cross-language variability with regard to the specific effects of smiling on formant frequencies, and therefore may also account for the differences in exact patterns of formant raising between our two studies.

Similar to studies of Norwegian and Dutch ID speech (Benders, 2013; Englund & Behne, 2005, 2006) we found differences in the distribution of F1 and F2 frequencies between ID and AD speech, but no overall increase in vowel space area. Moreover, our results agree with recent studies

(e.g., Benders, 2013; Cristia & Seidl, 2013; Lam & Kitamura, 2012) that across-the-board hyperarticulation is not generalizable to all aspects and contexts of ID speech. It is possible that the exclusion of tokens with high F0 may have affected the results of the current study by excluding the most hyperarticulated tokens (Adriaans & Swingley, 2012), but the similarity to other studies that found no difference between ID and AD vowel space (e.g., Englund & Behne, 2006) as well as the differences in formant frequencies found between speech styles make this explanation unlikely. The current study provides nuance to the formulation of a holistic approach to ID speech research that takes into consideration other possible phonetic, communicational, and situational factors, as suggested by Cristia and Seidl (2013; see also Englund & Behne, 2006; Lam & Kitamura, 2012). More research is needed to determine the conditions of systematic changes in the distributions of vowels as a function of different speech styles, speaking conditions, and materials, as well as what, if any, role these changes may play in infant language acquisition.

In addition to measurements of vowel space area, the current study employed measurements of vowel space dispersion as a method of quantifying the clarity of point vowels in ID as compared with AD speech (Bradlow et al., 1996; Liu et al., 2003). However, we found no difference in clarity between ID and AD speech. These results agree

with recent findings demonstrating that vowel categories are not necessarily more distinct in ID than in AD speech (Cristia & Seidl, 2013; McMurray, Kovack-Lesh, Goodwin, & McEchron, 2013). One possible cause of the lack of difference in vowel space dispersion between ID and AD speech could be a large within-category variance in ID vowel formant frequencies (Cristia & Seidl, 2013; Kuhl et al., 1997; McMurray et al., 2013). Kuhl et al. (1997) hypothesized that increased variability in ID speech may aid in language learning by allowing infants to attend to non-frequency-specific characteristics of vowel categories, rather than to the particular frequencies their mothers produce. However, McMurray et al. (2013) found that increased variability in ID vowel production caused a statistical learning model to perform better on vowel discrimination in AD rather than ID speech. This finding suggests that increased variability in ID speech may be a byproduct of other aspects of ID speech modifications, rather than a didactic tool (see also Benders, 2013). As Cristia and Seidl (2013) suggested, it is necessary to design further studies to assess the degree of variability across different sound categories in ID speech and whether this variability has any effect on infant speech processing.

The current study did not show any changes over time in the distribution of F1 and F2 frequencies or acoustic vowel space area for either a longitudinal (Study 1) or a cross-sectional (Study 2) design. This lack of a change suggests that vowel space area in maternal speech to infants is not significantly different compared with that to adults over the first 2 years of life, at least in certain interactive contexts (e.g., reading activities). These findings agree with previous research demonstrating that little to no change in vowel space areas is found in speech to children between 4 and 11 months in American English (Cristia & Seidl, 2013), 1 and 6 months in Norwegian (Englund & Behne, 2006), across the first year in Cantonese (Xu Rattanasone et al., 2013), or at 7 months and 5 years in Mandarin (Liu et al., 2009). These findings thus provide converging evidence that infants experience a vowel space that is consistent over substantial periods of time during early childhood. However, some researchers suggest that modifications to ID speech may be most pronounced around the time that the infant is first learning to speak (Bernstein Ratner, 1984, 1987; Ko, 2012). Although Study 2 examined speech to infants at 13 and 20 months of age, there was a large amount of time between these two visits, and the infants' stage of linguistic development was not assessed. Future research should therefore take a more temporally fine-grained approach to examining the characteristics of vowels in speech to older infants as a function of the infants' stage of linguistic development.

The results of the present study contribute to the recent findings suggesting that the modification of point vowels in ID speech to more distinctive positions relative to AD speech may depend on a number of communicative and situational factors (Cristia & Seidl, 2013; Englund & Behne, 2006; Kuhl et al., 1997; Lam & Kitamura, 2010, 2012). The present study used storybook materials, whereas many previous studies employed spontaneous or semispontaneous

speech (Burnham et al., 2002; Cristia & Seidl, 2013; Englund & Behne, 2005, 2006; Kondaurova et al., 2012; Kuhl et al., 1997; Liu et al., 2003; Liu et al., 2009; Uther et al., 2007; Xu Rattanasone et al., 2013). The use of a storybook allowed us to avoid the problem of linguistic context variability (e.g., consonantal environment, position of target words in an utterance, utterance length), which was not completely controlled in previous research but could affect vowel characteristics (Bernstein Ratner, 1986; Englund & Behne, 2005; Hillenbrand, Clark, & Nearey, 2001; Stevens & House, 1963). The use of a storybook also made it unlikely that we would need to exclude participants due to small numbers of tokens, which has been a problem with previous studies (e.g., Cristia, 2010). Finally, by using storybook speech, our study opens the door to the possibility of a storybook method for analyzing vowel formant characteristics of individual mothers in a clinical setting. The amount of linguistic control afforded by a storybook text allows for the possibility of an automated system that would require little work on the part of the clinician. Given evidence of an association between mothers' vowel space modifications and their infants' later speech discrimination skills, such a test may serve as a predictor of children's speech outcomes (Liu et al., 2003). Future work is necessary to investigate correlations between ID storybook speech and infant language acquisition, as well as to develop a specific clinical test of vowel production.

The storybook materials used here were sensitive enough to elicit reliable, significant differences in formant frequency shifts in ID compared with AD speech. Several studies examining spectral, acoustic, phonetic, and perceptual characteristics of storybook speech in ID and AD conditions found reliable differences between these two conditions (Inoue, Nakagawa, Kondou, Koga, & Shinohara, 2011; Jacobson, Boersma, Fields, & Olson, 1983; McMurray et al., 2013). Moreover, prosodic measures from speech samples in Study 2 support the assumption that the storybook method successfully elicited the ID speech register (Dilley, Burnham, Wieland, Kondaurova, & Bergeson, in preparation), rather than a form of simulated ID speech. For example, mean F0 was 28–43 Hz higher for ID than AD speech at each interval. Although not as pronounced a difference as Fernald et al. (1989) found (i.e., 102 Hz higher for ID than for AD speech), the differences in mean F0 for the present study were very similar to findings in Fernald and Simon (1984; i.e., 54 Hz higher for ID than for AD speech) and between ID and “baseline” speech in Jacobson et al. (1983; i.e., 35–45 Hz higher for ID than for AD speech). In contrast, Fernald and Simon (1984) found that simulated ID speech did not differ in mean F0 from AD speech, and Jacobson et al. (1983) found that simulated ID speech differed from baseline speech in mean F0 by only 13 Hz. Similarly, the speech styles in the current study also differed in maximum F0 and articulation rate, in a manner consistent with previous research (Fernald & Simon, 1984). Although there was no significant difference in utterance duration between speech styles, this apparent inconsistency with previous research (Fernald, 1989; Fernald & Simon, 1984)

may be attributed to the difficulty in manipulating utterance length in a storybook reading task. Overall, the speech samples used in the current study showed prosodic characteristics consistent with previous measurements of ID speech and underscore the validity of the storybook speech as examples of ID and AD speech constructs, respectively.

Given that the prosodic analyses of these speech samples showed differences consistent with an ID–AD contrast, including wider F0 range, slower rate, and shorter utterances for ID than AD speech, it is important to inquire why no increased vowel space expansion in ID speech was found here, whereas vowel space expansion was found in some (but not all) other studies. One possibility is that storybook speech may be relatively careful overall in both ID and AD speech conditions. This carefulness could lead to a kind of “ceiling effect” and inability to detect a reliable difference between the conditions. Another possibility is that the use of the storybook materials may distract mothers from being fully engaged with their infants, thus promoting low infant responsiveness. Lam and Kitamura (2012) proposed that vowel hyperarticulation may be affected by a reduced level of infant interactivity. In a study of 6-month-olds, the authors demonstrated that when an infant could fully or partly hear their mother’s voice the acoustic vowel space was expanded in ID relative to AD speech, but this expansion did not occur when the infant could not hear the mother’s voice. The authors hypothesized that when an infant is not able to hear the mother’s voice, mothers unconsciously decrease the distinctiveness between vowel categories and shift their attention to speech characteristics such as pitch that are known to elicit and maintain infant attention (Lam & Kitamura, 2010, 2012).

Recent research on 18-month-old infant and parent eye-gaze during a word learning task also demonstrated that the most successful learning occurred during the moments when the attention was coordinated between the infant and the parent (Yu & Smith, 2012). Analysis of infant and parent interactions suggested that parents provided object names at optimal moments when they were following their infant’s lead and interest in the attended object (Bornstein, Tamis-LeMonda, Hahn, & Haynes, 2008; Gros-Louis, West, Goldstein, & King, 2006; Miller, Ables, King, & West, 2009). Thus, it is possible that the lack of the expansion of acoustic vowel space in ID relative to AD speech in the present study could be accounted for by the use of the storybook materials affecting infants’ responsiveness and modulated attention. If mothers were not attending to their infants or following their responses while reading a storybook, there would be no need to promote language learning through exaggeration of vowel characteristics in the present study. These issues will need to be investigated in further research.

In summary, we found that although systematic differences existed in the distribution of vowel formant frequencies in ID as compared with AD speech, no expansion of the acoustic vowel space or greater clarity of vowels was observed across speech styles for storybook speech. Our results could be accounted for by our use of storybook

speech, which allowed novel control over a number of variables (Burnham et al., 2002; Cristia & Seidl, 2013; Englund & Behne, 2006; Kondaurova et al., 2012; Kuhl et al., 1997; Liu et al., 2003; Liu et al., 2009; Uther et al., 2007). Further studies are necessary to investigate the nature of modifications in ID speech, as well as how and whether this speech style promotes infant language acquisition (c.f., Cristia & Seidl, 2013). Overall, the present study reveals that expansion of the vowel space reported previously (e.g., Burnham et al., 2002; Kuhl et al., 1997) did not generalize to a different interactive context, namely, a storybook reading task. These results suggest that the presence and/or degree of hyperarticulation in ID speech is affected by factors other than speech style (infant or adult directed) and age of the addressee—specifically, the context of interaction matters.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grant R01 DC 008581, awarded to T. R. Bergeson. We thank members of the RAP Lab at Bowling Green State University and the MSU Speech-Perception Production Lab at Michigan State University for helping in the analysis of the speech tokens. We also would like to thank Shannon Aronjo, Erin Crask, Kabreea Dunn, Carrie Hansel, Heidi Neuburger, Brittne Ostler, Crystal Spann, Julie Wescliff, Heather Winegard, and Neil Wright for their help in preparing materials and recording and analyzing mothers’ speech.

References

- Adriaans, F., & Swingle, D. (2012). *Distributional learning of vowel categories is supported by prosody in infant-directed speech*. Paper presented at the 34th Annual Conference of the Cognitive Science Society, Sapporo, Japan.
- Baker, M. (1885). A collection of formulae for the area of a plane triangle. *Annals of Mathematics*, 2(1), 11–18.
- Benders, T. (2013). Mommy is only happy! Dutch mothers’ realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development*, 36, 847–862.
- Bernstein Ratner, N. (1984). Patterns of vowel modification in mother–child speech. *Journal of Child Language*, 11, 557–578. doi:10.1017/S030500090000595X
- Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother–child speech. *Journal of Phonetics*, 14, 303–309.
- Bernstein Ratner, N. (1987). The phonology of parent–child speech. In K. E. Nelson & A. Van Kleeck (Eds.), *Children’s Language* (Vol. 6, pp. 159–174). Hillsdale, NJ: Erlbaum.
- Boersma, P., & Weenink, D. (2012). Praat: Doing phonetics by computer (Version 4.0.26). [Computer software]. Available from <http://www.praat.org>
- Bornstein, M. H., Tamis-LeMonda, C. S., Hahn, C.-S., & Haynes, O. M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, 44(3), 867–874.
- Bowman, B. T., Donovan, M. S., & Burns, M. S. (Eds.). (2000). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academies Press.

- Bradlow, A. R., Torretta, G. M., & Pisoni, D. B.** (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20*, 255–272.
- Burnham, D., Kitamura, C., & Vollmer-Conna, U.** (2002, May 24). What's new, pussycat? On talking to babies and animals. *Science, 296*(5572), 1435–1435.
- Cristia, A.** (2010). Phonetic enhancement of sibilants in infant-directed speech. *Journal of the Acoustical Society of America, 128*, 424–434.
- Cristia, A., & Seidl, A.** (2013). The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language, 41*, 913–934. doi:10.1017/S0305000912000669
- Dilley, L. C., Burnham, E. B., Wieland, E. A., Kondaurova, M. V., & Bergeson, T. R.** (in preparation). *Bootstraps for word learning: Linguistic-prosodic variation in maternal speech to infants during the first two years*. Manuscript in preparation.
- Englund, K., & Behne, D. M.** (2005). Infant directed speech in natural interaction—Norwegian vowel quantity and quality. *Journal of Psycholinguistic Research, 34*, 259–280.
- Englund, K., & Behne, D. M.** (2006). Changes in infant directed speech in the first six months. *Infant and Child Development, 15*, 139–160. doi:10.1002/icd.445
- Evans, M. A., & Saint-Aubin, J.** (2005). What children are looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychological Science, 16*, 913–920.
- Fagel, S.** (2009). *Effects of smiled speech on lips, larynx and acoustics*. Paper presented at the International Conference on Auditory-Visual and Speech Processing, Norwich, United Kingdom.
- Fant, G.** (1973). *Speech sounds and features*. Cambridge, MA: MIT Press.
- Fernald, A.** (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60*, 1497–1510.
- Fernald, A., & Simon, T.** (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*, 104–113. doi:10.1037/0012-1649.20.1.104
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I.** (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language, 16*, 477–501.
- Fitzgerald, J., Spiegel, D. L., & Cunningham, J. W.** (1991). The relationship between parental literacy level and perceptions of emergent literacy. *Journal of Reading Behavior, 23*, 191–213.
- Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P.** (2006). Mothers provide differential feedback to infants' pre-linguistic sounds. *International Journal of Behavioral Development, 30*, 509–516.
- Hayashi, A., Tamekawa, Y., & Kiritani, S.** (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language, and Hearing Research, 44*, 1189–1200.
- Hillenbrand, J. M., Clark, M. J., & Nearey, T. M.** (2001). Effects of consonantal environment on vowel formant patterns. *Journal of the Acoustical Society of America, 109*, 748–763.
- Inoue, T., Nakagawa, R., Kondou, M., Koga, T., & Shinohara, K.** (2011). Discrimination between mothers' infant- and adult-directed speech using hidden Markov models. *Neuroscience Research, 70*, 62–70.
- Jacobson, J. T., Boersma, D. C., Fields, R. B., & Olson, K. L.** (1983). Paralinguistic features of adult speech to infants and small children. *Child Development, 54*, 436–442.
- Kent, R. D., & Read, C.** (1992). *The acoustic analysis of speech*. San Diego, CA: Singular.
- Kienast, M., & Sendlmeier, W. F.** (2000). *Acoustical analysis of spectral and temporal changes in emotional speech*. Paper presented at the ISCA Tutorial and Research Workshop on Speech and Emotion, Newcastle, Northern Ireland, United Kingdom.
- Kitamura, C., & Burnham, D.** (1998). The infant's response to maternal vocal affect. In C. Rovee-Collier, L. Lipsitt, & H. Hayne (Eds.), *Advances in infancy research*. (Vol. 12, pp. 221–236). Stamford, CT: Ablex.
- Kitamura, C., & Burnham, D.** (2003). Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy, 4*, 85–110.
- Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S.** (2002). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior & Development, 24*, 372–392.
- Ko, E.-S.** (2012). Nonlinear development of speaking rate in child-directed speech. *Lingua, 122*, 841–857.
- Kondaurova, M. V., Bergeson, T. R., & Dilley, L. C.** (2012). Effects of deafness on acoustic characteristics of American English tense/lax vowels in maternal speech to infants. *Journal of the Acoustical Society of America, 132*, 1039–1049.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., . . . Lacerda, F.** (1997, August 1). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*, 684–686.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T.** (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society, 363*, 979–1000.
- Labov, W., Ash, S., & Boberg, C.** (2006). *The atlas of North American English: Phonetics, phonology, sound change: A multimedia reference tool*. Berlin, Germany: Walter de Gruyter.
- Lam, C., & Kitamura, C.** (2010). Maternal interactions with a hearing and hearing-impaired twin: Similarities and differences in speech input, interaction quality, and word production. *Journal of Speech, Language, & Hearing Research, 53*, 543–555.
- Lam, C., & Kitamura, C.** (2012). Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation. *Developmental Science, 15*, 212–221.
- Liu, H.-M., Kuhl, P. K., & Tsao, F.-M.** (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science, 6*, F1–F10. doi:10.1111/1467-7687.00275
- Liu, H.-M., Tsao, F.-M., & Kuhl, P. K.** (2009). Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: A longitudinal study. *Journal of Child Language, 36*, 909–922.
- MathWorks.** (2009). MATLAB [Computer software]. Natick, MA: Author.
- McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W.** (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition, 129*, 362–378.
- Miller, J. L., Ables, E. M., King, A. P., & West, M. J.** (2009). Different patterns of contingent stimulation differentially affect attention span in prelinguistic infants. *Infant Behavior & Development, 32*, 254–261.
- Morrison, G. S., & Nearey, T. M.** (2011). FormantMeasurer: Software for efficient human-supervised measurement of formant trajectories [Computer software]. Available from <http://geoff-morrison.net>

- Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research, 51*, 574–585.
- Pitt, M., Dilley, L. C., Johnson, K., Kiesling, S., Raymond, W., Hume, E., . . . Fosler-Lussier, E. (2007). *The Buckeye speech corpus*. Retrieved from The Ohio State University Department of Psychology website: <http://www.buckeyecorpus.osu.edu>
- Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy, 13*, 410–420.
- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academies Press.
- Stern, D. N., Spieker, S. S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language, 10*, 1–15.
- Stevens, K. N. (2000). *Acoustic phonetics*. Cambridge, MA: MIT Press.
- Stevens, K. N., & House, A. S. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech, Language, and Hearing Research, 6*, 111–128.
- Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Perception and Psychophysics, 27*, 24–27.
- Tartter, V. C., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *Journal of the Acoustical Society of America, 96*, 2101–2107.
- Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication, 49*, 2–7.
- Vallabha, G. K., & Tuller, B. (2002). Systematic errors in the formant analysis of steady-state vowels. *Speech Communication, 38*, 141–160.
- Weisstein, E. (2015). *Triangle Area*. Retrieved from <http://mathworld.wolfram.com/TriangleArea.html>
- Xu, N., Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2013). Vowel hyperarticulation in parrot-, dog-, and infant-directed speech. *Anthrozoös, 26*, 373–380.
- Xu Rattanasone, N., Burnham, D., & Reilly, R. G. (2013). Tone and vowel enhancement in Cantonese infant-directed speech at 3, 6, 9, and 12 months of age. *Journal of Phonetics, 41*, 332–343.
- Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition, 125*, 244–262.

Appendix A

“Look What I Found” by Brittanie and Heather

(Tokens of interest are underlined)

The sweet pink kitten went for a walk and saw the cool green turtle. The cool green turtle found a little green key. Who did it belong to? The cool green turtle wanted the sweet pink kitten to help in finding who the key belonged to. As they were walking, the sweet pink kitten saw a small green ball. The sweet pink kitten and the cool green turtle were not sure who it belonged to. They picked up the small green ball and the little green key and kept walking. Along the way, the cool green turtle found a pretty blue crystal. Once again, the sweet pink kitten and the cool green turtle wanted to know who the pretty blue crystal belonged to. They picked up the pretty blue crystal along with the little green key and the small green ball and kept walking. Then they saw the cute brown dog. He looked very sad! The cute brown dog said, “Oh no! I have lost my little green key, my small green ball and my pretty blue crystal. I dropped them and cannot find them anywhere!” The sweet pink kitten and the cool green turtle were very happy that they found who the little green key, small green ball and pretty blue crystal belonged to. The cute brown dog wanted his things returned. The sweet pink kitten and the cool green turtle were glad to return them, and this made the cute brown dog very happy.

Appendix B

Vowel Space Dispersion Calculations

First, the centroid (C) of each speaker-condition vowel space triangle was calculated using the formula:

$$C = \left(\frac{F1_a + F1_i + F1_u}{3}, \frac{F2_a + F2_i + F2_u}{3} \right), \quad (A1)$$

where /i/, /a/, and /u/ were the corners of each vowel space triangle and $F1$ and $F2$ were the x and y coordinates of each of the corners.

Next, the Euclidean distance ($|d|$) of each token from the centroid was calculated using the formula:

$$|d| = \sqrt{(F1_C - F1_t)^2 + (F2_C - F2_t)^2}, \quad (A2)$$

where $F1_C$ and $F2_C$ were the x and y coordinates of the centroid and $F1_t$ and $F2_t$ were the first and second formant values for the token in question.

Finally, the vowel space dispersion (D) was calculated as the ratio of the Euclidean distances ($|d|$) of each token from the centroid of the triangle to the number of tokens (n), using the formula:

$$D = \frac{\sum |d|}{n}. \quad (A3)$$
