2010

LIP MOVEMENT EXAGGERATIONS DURING INFANT DIRECTED SPEECH

Jordan R. Green
University of Nebraska-Lincoln, jgreen4@unl.edu

Ignatius S. B. Nip
San Diego State University

Erin M. Wilson
University of Wisconsin-Madison

Antje Mefferd
Wichita State University, amefferd@unlserve.unl.edu

Yana Yunusova
University of Toronto, Faculty of Law, yana.yunusova@utoronto.ca

JSLHR Papers in Press. Published August 10, 2010, as doi: 10.1044/1092-4388(2010/09-0005)
The final version is at http://jslhr.asha.org.

Running head: **Lip Movement Exaggerations**

LIP MOVEMENT EXAGGERATIONS DURING INFANT DIRECTED SPEECH

Jordan R. Green[a], Ignatius S.B. Nip[b], Erin M. Wilson[c], Antje S. Mefferd[d] , &Yana Yunusova[e]

[a]Department of Special Education and Communication Disorders, University of Nebraska-Lincoln

[b]School of Speech, Language, & Hearing Sciences, San Diego State University

[c]Waisman Center, University of Wisconsin-Madison

[d] Communication Sciences and Disorders, Wichita State University

[e] Department of Speech-Language Pathology at the University of Toronto

Address editorial correspondence to:
Jordan R. Green, Ph.D.
jgreen4@unl.edu
Department of Special Education and Communication Disorders
University of Nebraska-Lincoln
318 Barkley Center
Lincoln, NE 68583

Abstract

Purpose: Although a growing body of literature has indentified the positive effects of visual speech on speech and language learning, oral movements of infant directed speech have rarely been studied. This investigation used 3-dimensional motion capture technology to describe how mothers modify their lip movements when talking to their infants.

Method: Lip movements were recorded from twenty-five mothers as they spoke to their infants and other adults. Lip shapes were analyzed for differences across speaking conditions. The maximum fundamental frequency, duration, acoustic intensity, and first and second formant frequency of each vowel were also measured.

Results: Lip movements were significantly larger during infant directed speech than during adult directed speech, although the exaggerations were vowel specific. All of the vowels produced during infant directed speech were characterized by an elevated vocal pitch and a slowed speaking rate when compared to vowels produced during adult directed speech.

Conclusion: The pattern of lip shape exaggerations did not provide support for the hypothesis that mothers produce exemplar visual models of vowels during infant directed speech. Future work is required to determine if the observed increases in vertical lip aperture engender visual and acoustic enhancements that facilitate the early learning of speech.

Lip Movement Exaggerations during Infant Directed Speech

When interacting with infants, adults significantly modify their communication style through changes to language, speech, and gesture. These adaptations may have a number of positive influences on early development, including the facilitation of child-parent bonding, attention, and affect control, and speech and language development (Bernstein Ratner, 1986; Dominey & Dodane, 2004; Fernald et al., 1989; Kuhl et al., 1997). Identifying both the auditory and visual features of IDS is an essential step toward understanding the contribution of environmental stimulation to the development of speech and language.

IDS is conveyed *acoustically* through voice and speech patterns and *visually* through facial, head, and body movements. The acoustic features of IDS are well-established. Relative to adult directed speech (ADS), IDS is characterized by a slowing of rate, an increase in pause frequency and duration, an increase in the mean and range of fundamental frequency (Amano, Nakatani, & Kondo, 2006; Fernald & Simon, 1984; Grieser & Kuhl, 1988; Katz, Cohn, & Moore, 1996; Stern, Spieker, Barnett, & MacKain, 1983; Swanson, Leonard, & Gandour, 1992), as well as an increase in the acoustic distance between vowels (Kuhl et al., 1997). Although typically-developing children use both auditory and visual information for learning to comprehend and produce spoken language, the characteristics of the visual component of IDS, such as facial and lip movements, have not been identified.

Facial motion during speech provides infants with a rich and salient source of speech and language cues. Research on multimodal perception of speech in adults has clearly demonstrated that watching a speaker’s facial and head movements markedly

improves the speed and accuracy of sound discrimination as well as auditory comprehension, particularly in noisy environments (e.g., Bernstein, Takayanagi, & Auer, 2004; Grant & Seitz, 2000; MacLeod & Summerfield, 1987; Munhall, Jones, Callan, Kuratate, & Vatikiotis-Bateson, 2004; Rosenblum, Johnson, & Saldaña, 1996; Sumby & Pollack, 1954; van Wassenhove, Grant, & Poeppel, 2005). The improved processing afforded by visible speech can be explained, in part, by the complementary and redundant cues for place and manner of articulation, duration, prosody, rhythmicity, and intensity (e.g., Cho, 2005; Edwards, Beckman, & Fletcher, 1991; Erickson, 1998; Summerfield, 1987; Summers, 1987).

The mouth may be a particularly potent visual stimulus for gaining an infant's attention and then providing complimentary and redundant speech cues. In early speech learning, watching the mouth is also important for establishing links between the visual and acoustic representations of speech sounds and, potentially, for learning speech through imitation (Kuhl & Meltzoff, 1982; Patterson & Werker, 2003). For example, research on young infants has demonstrated infants learn to recognize the oral postures that accompany different vowels, prior to the onset of well-developed vocalizations (Kuhl & Meltzoff, 1982; Patterson & Werker, 2003). These early associations between facial and speech combinations may be a consequence of infants' strong attentional bias towards moving faces (Biringen, 1987; Cohn & Elmore, 1988; Slater & Kirby, 1998; Toda & Fogel, 1993), and their precocious ability to learn associations between moving objects seen and synchronously heard (Gogate & Bahrick, 2001).

One relatively unexplored possibility is that parents exaggerate the visible aspects of speech to facilitate early speech and language learning. Findings from studies of hand

and body gestures during IDS suggest that gestures directed toward infants are timed to direct the child's attention to important cues regarding the structure of language (Brand, Baldwin, & Ashburn, 2002). Exaggerated oral movements may serve a similar function in early development.

Of particular interest to the current investigation is the possibility that parents exaggerate their oral movements to enhance the distinction among vowels. Several studies on the acoustic characteristics of IDS have reported that parents maximized the acoustic distinction among different vowels when speaking to their infants (Burnham, Kitamura, & Vollmer-Conna, 2002; de Boer, 2003; Kuhl et al., 1997). Such exaggerations in speech are referred to as hyperarticulations (Lindblom, 1990) and are known to enhance speech clarity and intelligibility (Payton, Uchanski, & Braida, 1994; Picheny, Durlach, & Braida, 1985; Smiljanic & Bradlow, 2005). It is not known if parents similarly hyperarticulate their oral movements to enhance the visual distinction among vowels. As displayed in Figure 1, parents could easily hyperarticulate vowel sounds in the visual domain by, for example, increasing the lip opening for open vowels such as "a" and the lip spread for "ee," and by exaggerating lip rounding for "oo."

insert Figure 1 about here

This investigation uses 3-dimensional (3D) motion capture technology to describe how mothers modify their articulatory movements for vowels when communicating with their infants. The following four experimental questions will be addressed: (1) Do mothers exaggerate articulatory movements during IDS relative to how they articulate during ADS? (2) Are lip shapes for vowels more distinctive during IDS than during ADS? (3) Are there individual differences among mothers in the degree of articulatory

exaggeration? and (4) Across mothers, is there an association between the degree of exaggeration in the facial movement and acoustic characteristics of IDS?

Method

Participants

The participants were 25 English-speaking mother-infant dyads who were enrolled in a study on early speech motor development. There were 25 mothers between 23 and 42 years of age (M = 32.9, SD = 5.32), and 12 male and 13 female infants. Post secondary maternal education levels ranged between 0 to 10 years (M = 4.1, SD = 2.3). The data from two additional mothers was not included because they tended to look downward, which placed their face out of the cameras' fields of view. All mothers had negative histories of neurologic or severe visual impairment and did not show any evidence of speech, language, or voice disorders.

All infants were between nine and ten months of age. This age was selected because mothers are likely to be very active in modeling articulations at this stage of development when their children are learning to produce sounds and recognize words. More specifically, by the second half of the first year, infants are in the early stages of learning to understand words (e.g., Fenson et al., 1994) and to produce vowels in babble (Kent & Murray, 1982; Robb, Chen, & Gilbert, 1997; Rvachew, Slawinski, Williams, & Green, 1996). Moreover, prior studies, show that acoustic features of IDS, such as increased F0 (Amano et al., 2006), are produced throughout the first year of a child's life. Because infant directed speech may change depending on the infants developmental status (Englund & Behne, 2006), future work will need to investigate infants at different

stages of development, particularly during the first half of infancy when the foundations of speech perception are established.

All children were from monolingual, English speaking homes. All infants had negative histories of neurologic or visual impairment. All infants performed at age level on the Battelle Developmental Inventory, 2nd edition (Newborg, 2005) when they returned for a follow-up visit at 12-months of age. On the day of data collection, all mothers passed a binaural 25 dB pure-tone screening at 1000, 2000, and 4000 Hz.

Speaking Conditions

Facial movements were recorded from each mother during four different speaking conditions: (1) storytelling to infant, (2) storytelling to adult, (3) story reading to infant, and (4) story reading to adult. Both tasks were designed to elicit the target words: “beet,” “bat,” “boot,” and “Bobby.” These words were chosen because their medial vowels have well-defined acoustic and visual targets that circumscribe the boundaries of vowel space (see Figure 1). For the story production task, mothers were provided with an illustrated book without words and were instructed to tell the story depicted in the pictures about a little boy named Bobby and his experiences playing baseball and picking beets in the garden. For the reading task, mothers were given the same illustrated book with words designed for young children. The storytelling and reading tasks were intended to provide two different contexts for eliciting IDS, with the former less constrained than the latter. The order of tasks was not varied so that the mothers were consistently producing a story rather than retelling a story. Because of the story telling task, the number of repetitions of each word varied across subjects (M = 9; SD = 3.4).

During the adult directed speech (ADS) task, baseline measures of articulatory movements and speech were recorded while each subject read and told the story to a laboratory assistant. During the IDS task, mothers were instructed to perform the same tasks while speaking to their babies as they typically do. For the IDS tasks, mothers were positioned facing the child, who was either secured in an infant-seat approximately two feet in front of the mother or in several cases, sitting on the mother's lap. In these cases, the infants were positioned in a way that permitted them to view both their mother's face and the book. To minimize self-consciousness on the part of the mother, the motion capture cameras were partly camouflaged by a wall-sized mural that depicted a jungle scene (see Figure 2). In addition, mothers and infants were separated from the investigators by a black curtain that surrounded the data collection area.

insert Figure 2 about here

Audio and Lip Movement Recordings

Digital audio recordings (Fs = 44.l k Hz, 16 bit linear PCM) were made through the entire session using a professional quality lapel microphone that was mounted on each mother's forehead. The microphone was head-mounted to ensure that the mic-to-mouth distance was kept constant during the entire data collection session. Movements of the lower and upper lip were captured in 3D at 120 frames per second using an 8-camera optical motion capture system (Eagle Digital System, Motion Analysis Corp). Prior to data collection, the system was calibrated according to manufacturer specifications.

The motion capture system tracked the movement of spherical reflective markers (approximately 2 mm in diameter) that were illuminated with an infrared light source. Fifteen markers were placed on the mother's face in the following regions: forehead,

eyebrow, nose, lips, and jaw. Only the upper lip (UL), lower lip (LL), right (RC), and left corner (LC) markers were studied. An example of the marker placement is displayed in Figure 3. The UL and LL markers were located midline on the vermillion border of each lip. The lower lip marker represented the combined motions of the jaw and lower lip. The RC and LC were located near the right and left of the oral commissures, just lateral to the labionasal fold.

insert Figure 3 about here

The 3D positional data from each marker was expressed relative to a room-based coordinate system. Following position tracking, the movement signals were digitally low-pass filtered (f_{lp} = 10 Hz) using a zero-phase shift forward and reverse digital filter (Butterworth, 8 pole). Two signals were derived from the 3D time-histories of each lip marker: (1) lip separation, which was defined by the 3D Euclidean distance between UL and LL, and (2) lip spread, which was defined by the by the 3D Euclidean distance between the markers at the mouth corners (i.e., RC and LC).

Speaking Rate

The speaking rate in words per minute (WPM) was computed for the entire book reading by dividing the number of words produced by the time in minutes that it took to read the book. This measure included all of the pauses in the speech sample. Occasional comments produced by the mothers during reading were not included in the calculation of speaking rate.

Acoustic Analyses

The maximum fundamental frequency (F_0) and duration of each vowel was measured to validate that the participants were producing IDS and to determine if the

mothers who exhibited IDS in the acoustic domain also exhibited exaggerated facial movements.

Vocal pitch. Increased vocal pitch is a well-established feature of IDS (Amano et al., 2006; Fernald et al., 1989, McRoberts & Best, 1997; Swanson et al., 1992). Fernald and Kuhl (1987) suggest that F_0 changes are particularly salient for conveying communicative intentions in IDS. In the current study, F_0 contours were obtained for the vowel in each target word using the autocorrelation algorithm in TF32 (Milenkovic, 2004), followed by occasional hand correction of mistrackings. The maximum F_0 value was calculated from each pitch contour.

Vowel duration. The duration of each vowel was determined acoustically as an indicator of speaking rate differences between IDS and ADS speech. Vowels were segmented based on a spectrographic display using TF32 (Milenkovic, 2004). Boundaries were determined visually by the onset and offset of acoustic energy associated with both voicing and formants. Estimates of acoustic boundaries were confirmed through audio playback.

Vocal RMS. The maximum RMS level (dB) was also obtained for the entire reading passage and each vowel segment using TF32 (Milenkovic, 2004). This measure was used to determine if IDS is also associated with increased speech intensity. Because articulatory movements and muscle activity are known to increase with increased speech intensity (Dromey & Ramig, 1998; Schulman, 1989; Wohlert & Hammen, 2000), changes in speech intensity may account for potential increases in articulatory displacement during IDS.

Kinematic Analyses of Extent of Articulatory Range of Motion During Storybook Reading

As a global measure of articulatory working space, two standard deviation (2SD) ellipsoids were fit around the 3D motion paths that were recorded during each mother's reading of the entire storybook. The ellipsoid algorithm calculated the two standard deviation boundaries along the derived principle axis of motion and two additional orthogonal axes. The volume (mm^3) defined by the 2SD ellipsoid served as measure articulatory working space, which was compared across the IDS and ADS conditions.

The lip kinematic data that was recorded during the entire reading passage were also used to quantify changes in mouth shape across speaking conditions. For this analysis, we measured the maximum vertical (i.e., the maximum 3D Euclidean distance between the UL and LL markers) and horizontal aperture (i.e., the maximum 3D Euclidean distance between the RC and LC markers) across the entire passage. These measures were taken, in addition to the working space measure, to determine specific changes in lip shape. Presumably, the maximum vertical and horizontal aperture would be sensitive to potential exaggerations of lip opening and spreading, respectively, during vowel opening. Although the features that best capture lip protrusion are multidimensional and not fully understood, a study by Fromkin (1964) on vowel lip shape suggest that horizontal lip aperture effectively encodes information about the rounding feature of vowels. This study also suggested that, across vowels, changes in horizontal lip aperture are moderately to strongly coupled with anterior-posterior lower lip movements. In the current study, maximum aperture was used, rather than other measures such as average or standard deviation of aperture, because this measure is less affected by the observed differences in speaking rate across speaking tasks. The duration

of the reading passages were also recorded to test for differences in speaking rate across conditions.

Kinematic Analyses of Vowel Specific Changes

As displayed in Figure 4, for each word, lip opening during the transition from the initial bilabial consonant to the medial vowel was analyzed for each word. Audio playback from the digital video, which was time aligned with the kinematic data, was used to identify the lip movements associated with each vowel. The onset of lip opening for the vowel was the minima in a signal that represented the 3D Euclidean distance between the upper and lower lip during bilabial closure. The offset of lip opening for the vowel was the point that was associated with the maximum 3D distance between the lips during the vowel. Once each opening gesture was identified, they were measured for maximum vertical (i.e., the maximum 3D Euclidean distance between the UL and LL markers) and horizontal aperture (i.e., the maximum 3D Euclidean distance between the RC and LC markers).

insert Figure 4 about here

These kinematic descriptors of lip shape were used to characterize potentially important visual features (i.e., lip spreading, rounding, and vertical separation) used to discriminate among some vowel categories. More specifically, if the mothers' goal was to produce lip-shapes with maximal visual cue contrast to their infants, they would be expected to exaggerate the specific features that distinguish each vowel. For example, an increase in vertical lip aperture would be expected during "bat" and "Bobby," an increase in horizontal aperture would be expected for "beet," and a decrease in lip opening would be expected during an exaggerated version of lip rounding for "boot." An exaggerated

"boot" may also be produced with a decrease in both vertical and horizontal aperture for rounding.

To determine if lip shapes for vowels became more distinctive during IDS than during ADS, the Euclidean distance between vowels in lip shape workspace (lip separation vs. lip spread) were computed. Lip shape workspace (see Figure 1) was defined by the maximum vertical distance between the upper and lower lip (vertical aperture) and the maximum horizontal distance between the corners of the lip (horizontal aperture) during vowel opening for each word.

Acoustic Analyses of Vowel Specific Changes

Measurements of the first (F1) and second formants (F2) associated were identified on a broad-band spectrogram (260 Hz) for each target vowel. The onset of the vowel was identified as the first glottal pulse following the release of the burst and the offset was the last glottal pulse apparent in F1. Pratt 5.1 software (Boersma & Weenink, 2009) was used to extract the F1 and F2 time-histories of each vowel using the recommended default setting for extracting formants from the speech of female talkers (i.e., maximum formant 5000 Hz, window length = 0.0.25 s, pre-emphasis = 50 Hz). All formant time-histories were reviewed visually prior to analysis. In the rare event of mistrackings (< 20 samples), the data analysts were instructed to obtain a different sample of the vowel from the recordings. The extracted formant trajectories were then imported into a custom program written for Matlab (2009) that calculated the maximum F1 and F2 values within the mid 80% section of the vowel. This mid section was isolated to minimize coarticulation effects from flanking consonants. The maximum frequency

value was used because it paralleled the kinematic measures (i.e., maximum vertical and horizontal) used to quantify lip opening during each vowel.

Measurement Reliability

Ninety acoustic recordings were randomly selected for re-analysis of maximum fundamental frequency, vowel duration, and maximum RMS intensity. The intrarater reliability across all three measures was very high (Cronbach $\alpha = .99$). The intrarater reliability of the speaking rate measure was determined by reanalyzing ten files. The difference between first and second measurement was 1.45 WPM. Reliability analysis was not performed on the measures of articulatory working space, maximum vertical aperture, maximum horizontal aperture, and formant frequencies because these measures were algorithmically identified.

Statistical Analyses

The effects of speaking condition (IDS vs. ADS) and vowel (/i, a, ae, u/) were analyzed using a two-way repeated measures ANOVA for each kinematic measure to test for potential changes in lip shape as a function of speaking task. Measures obtained across the multiple repetitions of each word were averaged for each subject. The Holm-Sidak method was used for testing all pairwise multiple comparisons when significant differences were found. An overall alpha of $p < 0.05$ was used for all statistical testing. Data from the story telling and reading tasks were collapsed because an ANOVA revealed no significant differences in lip separation and spread for these tasks.

Results

Speaking Rate

Speaking rate in WPM was slower during IDS (M = 107.11, SEM = 3.55) than during ADS (M = 122.53, SEM = 5.04), F (1, 24) = 14.11, p < 0.001.

F_0, Duration, and Intensity

Vowel maximum F_0 and duration during each speaking condition are displayed in the top panel and bottom panel of Figure 5, respectively. Maximum F_0 was 57.32 Hz higher during IDS than during ADS, F (1, 24) = 5.45, p < 0.05. The samples of the entire reading passage were on average 38 seconds longer in duration during IDS than during ADS, F (1, 18) = 34.395, p < 0.001. Single vowels were on average 25 ms longer in duration during IDS than during ADS, F(1, 24) = 4.82, p < 0.001. The maximum RMS intensity of the vowels was not significantly different across the speaking conditions for the entire reading passage or for single vowels.

insert Figure 5 about here

Extent of Lip Movement Across Tasks during Book Reading

The working spaces (mm^3), as measured by the 2 standard deviation ellipsoid, for all the mouth markers (see Figure 3) were significantly larger during IDS than during ADS. Post hoc analysis of task effects revealed significantly greater working spaces for all articulators during IDS than during ADS: mean difference for UL = 611.60 mm^3, p < 0.001; mean difference for LL = 1537.92 mm^3 , p < 0.003; mean difference for RC = 990.04 mm^3, p < 0.001; mean difference for LC = 611.59 mm^3, p < 0.003. Summary statistics for the values obtained for vertical and horizontal lip aperture for the entire reading passage and the target vowels are reported Table 1. No task effects were

observed for the maximum horizontal aperture measure; however, the maximum vertical aperture was significantly larger during IDS than during ADS, $q(1, 17) = 5.51$, $p < 0.05$).

insert Table 1

Vowel-Specific Changes in Orofacial Movements Across Tasks

Data from two participants were excluded from the ANOVA model because their data set was missing at least one condition due to motion tracking errors. Vertical lip aperture was significantly larger during IDS than during ADS, $F(1, 24) = 64.85$, $p < 0.001$). Differences between lip apertures for IDS and ADS are shown for each target word are shown in Figure 6. Mean differences between IDS and ADS are displayed for all the mothers' data combined and for a subset of the data representing the ten mothers who exhibited the greatest increase in vertical aperture. The data in Figure 6 suggest that regardless of the degree of exaggeration, participants tended to exaggerate low vowels more than high vowels. Post hoc comparisons of the data set containing data from all of the mothers revealed significant differences in vertical lip aperture for /ae/ in "bat" [$t(24) = 7.10$, $p < 0.05$], /a/ in "Bobby" [$t(24) = 5.53$, $p < 0.05$], and /i/ in "beet" [$t(24) = 2.44$, $p < 0.05$]. Horizontal lip aperture did not differ significantly between IDS and ADS for any vowels.

Insert Figure 6 about here

Statistically significant differences were observed across mothers in the extent to which they exaggerated their articulatory movements during IDS, $F(1, 24) = 18.94$, $p <$

0.001. Across subject differences in articulatory exaggerations during IDS were further examined for the movements associated with the word “bat.” This word was selected because it was associated with the greatest articulatory change across speaking conditions. In Figure 7, the difference between the vertical aperture during IDS and ADS is plotted as a percentage change of vertical aperture (from minimum separation during consonantal closure subtracted to the maximum separation during vowel opening) during the opening gesture for each vowel. Seven of the mothers exhibited a 4% percent or less change and the remaining 20 mothers fell along a continuum with a 47% increase representing the greatest exaggeration of vertical lip aperture during IDS.

The data from the ten mothers who exhibited the highest change in F_0 during IDS as compared to during ADS are displayed in Figure 6. These data were plotted separately to identify aspects of articulatory exaggeration that may have been obscured by the group data. The results based on these ten mothers’ data were similar to those that were based on the entire data set (i.e., all 25 mothers) with the increase in lip-shape distinctiveness among vowels during IDS primarily driven by an increase in vertical aperture of /a/ and /ae/.

insert Figure 7 about here

Vowel Formant Changes Across Tasks

Maximum F1 and F2 values for each target vowel during IDS and ADS are displayed in Figure 8. Maximum F1 values were significantly larger during IDS than during ADS for all the target vowels: /ae/ [$t(24) = 7.10$, $p < 0.05$], /a/ [$t(24) = 5.53$, $p < 0.05$], /i/ [$t(24) = 2.44$, $p < 0.05$], and /u/ [$t(24) = 7.10$, $p < 0.05$]. Maximum F2 values

were only significantly larger for /ae/ ([t (24) = 7.10, p <0.05]) and /a/ ([t (24) = 7.10, p <0.05]).

insert Figure 8 about here

Correlations Between Acoustic and Kinematic Variables

Analyses were performed to determine if the mothers who exaggerated in the kinematic domain also exaggerated acoustic aspects of speech. For each subject, the average difference between IDS and ADS was computed for vertical lip aperture (mm), maximum F0 (Hz), and duration (ms) for /ae/. The degree of vertical aperture was moderately associated with maximum F0 [r (25) = 0.45, p < 0.03], duration [r (25) = 0.53, p < 0.001], and maximum F1[r (21) = 0.44, p < 0.04], but not with maximum F2.

Distinction Among Vowel Lip-Shapes

To determine if mothers enhanced the distinction among vowel lip shapes during IDS, the Euclidean distance between all possible vowel pairs in lip-shape space was compared across speaking conditions. The location of each vowel in lip-shape space was determined by its horizontal (x axis) and vertical (y axis) aperture during maximum opening. Four out of the six vowel pairs were significantly farther apart in lip-shape space during IDS than during ADS: /ae/ - /a/ = 2.0 mm [t (24) = 2.0, p < 0.05]; /ae/-/u/ = 2.94 mm, [t (24) = 3.34, p < 0.05]; /ae/-/i / = 2.16 mm [t (24) = 2.37, p < 0.05]; and /a/-/u/ = 1.95 mm [t (24) = 2.15, p < 0.05]. These findings suggest that the increased vertical lip aperture during IDS primarily increased the distance between the low vowels (i.e., /ae/ and /a/) and between these vowels and the other two vowels (i.e., /i/ and /u/).

Discussion

Lip Movements Were Exaggerated During IDS

The primary purpose of this investigation was to determine if mothers modify their lip movements when communicating with their infants. We hypothesized that IDS would be produced with larger mouth openings than ADS. To examine vowel-specific changes in lip shape, the size of vertical and horizontal lip aperture during IDS and ADS were measured during the production of two low vowels (i.e., /a/, /ae/) and two high vowels (i.e., /i/, /u/), which were embedded in target words. The primary findings were that mouth opening was significantly larger during IDS than during ADS, and that vertical lip aperture for low vowels was larger during IDS than during ADS. In contrast to this kinematic-based finding, all of the vowels (i.e., high and low) produced during IDS were characterized by an elevated vocal pitch and a slowed speaking rate when compared to vowels produced during ADS. The findings for changes in F1 and F2 frequencies were consistent with the mouth opening data; compared to vowels produced during ADS, all vowels during IDS were produced with higher F1 frequencies and only /a/ and /ae/ were produced with higher F2 frequencies. The 25 mothers varied considerably in the degree of articulatory exaggeration with mothers who exhibited acoustic exaggerations of F0 also tending to be the ones to exhibit exaggerated lip movements.

Contrary to our hypothesis, articulatory exaggerations during IDS were only in the degree of vertical aperture and not in other features of lip-shape such as spread and possibly rounding, as indicated by the finding of no differences across the speaking conditions in horizontal aperture. The observed increase in vertical aperture did, however,

increase the distinction among the four vowels in lip-shape space. Important issues to consider are the auditory-visual consequences of articulatory exaggerations in IDS and their potential to facilitate early speech learning.

Mothers Varied in the Extent to Which They Exaggerated Their Lip Movements

The significant changes in fundamental frequency and speaking rate across speaking tasks suggests that mothers in this study produced IDS during the experimental task. Therefore, the experimental paradigm appeared to elicit IDS even though (1) the speaking tasks were more structured than during a typical child-mother interaction and (2) the mothers were required to remain seated during the entire data collections session. As suggested by the data in Figure 4, the extent of exaggeration varied considerably among mothers. The factors that determine the extent to which a given mother produces IDS are not fully known, but potentially include the mother's and child's mood and personality, the infant's responsiveness, and the familial or ethnic culture. In addition, in the current study and others, it is possible that some mothers' willingness to produce IDS was suppressed by their awareness of being observed.

Speculation About Enhancements of Articulatory Exaggerations to Early Speech Learning

Despite its apparent lack of specificity with regard to visible features of vowels, the exaggerated lip opening observed during IDS could produce visual and acoustic enhancements that facilitate the early learning of speech.

Potential visual enhancements. Increasing lip opening may be a simple, effective strategy for focusing the child's attention to the face. Many studies have shown that children have a strong attentional bias toward moving faces over still faces (Biringen,

1987; Cohn & Elmore, 1988; Toda & Fogel, 1993). This effect is so robust it has been labeled the “still-face” effect (see Adamson & Frick, 2003, for review).

Exaggerated lip openings may also convey enhanced articulatory cues for vowels by modifying the luminance of the mouth and increase the visibility of the teeth and tongue (Erber, 1974; Rosenblum et al., 1996; Summerfield, MacLeod, McGrath, & Brooke, 1989). A study on the articulatory movements during the consonant /b/ suggest that the best visual exemplars of this sound, as judged by adults, are ones that are produced with greater lower lip displacements and speeds (Hall, Green, Moore, & Kuhl, 1999). Vowels produced during IDS may similarly have time-varying visual cues that make them good exemplars. For example, the slowing of speech during IDS may convey important duration cues that emphasize the distinction among vowels or even consonants (Klatt, 1976).

Why were the exaggerations only in the vertical aperture of low vowels? The absence of changes in horizontal aperture across speaking conditions might be interpreted to suggest that the mothers were not exaggerating to enhance specific sound contrasts (i.e., /i/ vs. /u/). Increased lip spreading and rounding (via changes to horizontal aperture) may be unnecessary because these sounds are already visually distinct (Montgomery & Jackson, 1983). In contrast, exaggerating the articulatory cues for /a/ and /ae/ may be productive because the lip shapes for these vowels are less distinguishable than are those of the other vowels (Montgomery & Jackson, 1983). Another possibility is that lip exaggerations are primarily implemented through increases in jaw opening with little contribution of lower lip movement. In this case, changes to lip shape would only be in the vertical aperture.

The absence of exaggerations in lip spread and protrusion could also be due to the insensitivity of the horizontal lip aperture measure to small changes in lip rounding and spreading. Lip protrusion, in particular, is multidimensional and the most sensitive measure of lip protrusion is currently unknown. An early study by Fromkin (1964) suggests that horizontal lip aperture and anterior-posterior lip movements are moderately to strongly coupled during the production of English vowels. Measures of lip movement along the anterior-posterior dimension have been used previously to quantify lip protrusion during speech (e.g., Goffman, Smith, Heisler, & Ho, 2008; Perkell, Matthies, Svirsky, & Jordan, 1993). The data from these studies suggest that some speakers only minimally protrude their lips during the production of /u/. For example, half of the subjects in Perkell and colleagues' (1993) study showed negligible lip protrusion in their investigation of the rounded vowel /u/ using the anterior-posterior lip movement measure. In addition, similar to the current study, the investigation by Goffman and colleagues (2008) showed only small differences (approximately 1 mm) in anterior-posterior lip protrusion between the words *beet* and *boot* for adult talkers. Additional studies are needed to determine the most sensitive measures of lip rounding and the speech contexts that elicit a strong rounding feature.

Even if the exaggerated vertical lip movements of IDS do not convey enhanced articulatory cues for specific sounds, they may convey prosodic information. Many studies have shown that visual prosody that is conveyed through head and face movements (Beckman, Edwards, & Fletcher, 1992; Erickson, 1998; Harrington, Fletcher, & Roberts, 1995; Summers, 1987) significantly facilitates the perception of speech (Munhall et al., 2004; Rosenblum et al., 1996). For infants, the strong marking of

prosody may be particularly effective for facilitating the perceptual segmentation among sounds, syllables, words, and phrases (Jusczyk, Hohne, & Mandel, 1995; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Wright Cassidy, 1989; Thiessen, Hill, & Saffran, 2005). Additional research is required to determine if the observed changes in lip movement during IDS are qualitatively similar to those used to mark stress.

Potential acoustic/auditory enhancements. The observation that mothers slow their speaking rate during IDS corroborates prior findings of increased vowel durations during IDS (Andruski & Kuhl, 1996; Bernstein Ratner & Luberoff, 1984; Fernald & Simon, 1984; Kuhl et al., 1997; Uther, Knoll & Burnham, 2007). The slowing of speech during IDS may not only afford extra processing time for the infant, but may also yield hyperarticulated acoustic vowel targets (Moon & Lindblom, 1994; Turner, Tjaden, & Weismer, 1995), which have been reported in prior studies of vowel spectral changes during IDS (Burnham et al., 2002; de Boer, 2003; Kuhl et al., 1997). Prior research has also shown that articulatory movement of vowels become slightly exaggerated when speech is slowed (Dromey & Ramig, 1998; Mefferd & Green, in press).

One possibility is that mothers may have exaggerated their lip openings for low vowels to produce the acoustic distinction among vowels during IDS, which has been reported previously in the literature (de Boer, 2003; Burnham et al., 2002; Kuhl et al., 1997). The observed increase in both F1 and F2 frequency is consistent with the findings of prior acoustic studies of IDS (Burnham et al., 2002; Kuhl et al., 1997; Liu, Kuhl, & Tsao, 2003). Consistent with some of these prior studies, was the observation of greater F1 change across conditions in low vowels than in high vowels (Burnham et al., 2002; Kuhl et al., 1997). This acoustic change associated with IDS is the expected consequence

of the articulatory change that was observed - an increase in vertical lip aperture (Lindblom & Sundberg, 1971; Stevens, 2000; Stevens & House, 1955).

One unexpected finding was that IDS was not produced with greater acoustic intensity than ADS. In general, larger lip openings are expected to transmit speech energy more efficiently than do smaller ones (Fairbanks, 1950). Moreover, many studies have shown that when talkers are asked to speak loudly they exaggerate their articulatory movements (Dromey & Ramig, 1998; Schulman, 1989; Tasko & McClean, 2004). Although speech loudness changes during IDS have rarely been investigated, a recent study of Jamaican talkers observed no significant speech intensity differences between IDS and a citation speaking task (Beckford Wassink, Wright, & Franklin, 2007). One perceptual study of synthesized IDS observed that infants are less interested in amplitude modulations than F_0 modulations (Fernald & Kuhl, 1987). Therefore, mothers may not exaggerate their speech intensity during IDS because their infants are not particularly responsive to such changes.

Motherese as an Affective Speaking Mode

The final consideration is that the articulatory changes observed during IDS may be a byproduct of facial expressions that are unique to and, perhaps, exaggerated during IDS (e.g., Chong, Werker, Russell, & Carroll, 2003). Mother's facial expressions are an important stimulus for affect attunement and engaging the child in meaningful social interaction (Kaplan, Bachorowski, Smoski, & Hudenko, 2002; Murray & Trevarthen, 1985; Stern, 1985; Trainor, Austin, & Desjardins, 2000). The heightened affect characteristic of IDS may also enhance children's motivation to communicate (Kitamura & Burnham, 1998; Locke, 1993) and may be especially salient to infants. Future work is

needed to elucidate how lip-shapes for different vowels are affected by overlaid facial expressions.

Study Limitations

One limitation of the current work is the restrictive laboratory setting in which the speech samples were collected. Outside the laboratory, mothers may exaggerate their speech to a greater extent than what was observed in this study and in ways that are qualitatively different from those observed in this laboratory study. In addition, although we positioned the infants in a way to maximize the likelihood that they would gaze at their mothers' faces, gazing patterns were not monitored. Mothers may be more inclined to exaggerate when their infants are actively staring at their mouths. Moreover, the method used to record mouth movement relied on the use of facial markers. Although the infants did not seem to be preoccupied with the markers, how their presence influenced the communication between infant and mother is uncertain. Finally, additional work is needed to determine if the observed, small changes in lip opening are perceptible to young infants.

Summary

Understanding the role of IDS in shaping early communication may have important implications for both theory and clinical practice. The current findings suggest that exaggerated lip movements are a characteristic of IDS, particularly during the production of low vowels. Mothers varied along a continuum in the extent to which they exaggerated their articulatory movements. The pattern of lip shape exaggerations did not provide strong support for the hypothesis that mothers were producing exemplar visual models of vowels during IDS. Additional research is required to understand the potential

significance of these exaggerations on communication development and to determine if the observed exaggerations are consistent with those observed in less restrictive data collection environments.

References

Adamson, L. B., & Frick, J. E. (2003). The still face: A history of a shared experimental paradigm. *Infancy, 4*, 451-473.

Amano, S., Nakatani, T., & Kondo, T. (2006). Fundamental frequency of infants' and parents' utterances in longitudinal recordings. *The Journal of the Acoustical Society of America, 119*, 1636-1647.

Andruski, J., & Kuhl, P. K. (1996). The acoustic structure of vowels in mothers' speech to infants and children. *Proceedings of the Fourth International Conference on Spoken Language Processing, Philadelphia, PA*, 1545-1548.

Beckford Wassink, A., Wright, R. A., & Franklin, A. D. (2007). Intraspeaker variability in vowel production: An investigation of motherese, hyperspeech, and Lombard speech in Jamaican speakers. *Journal of Phonetics, 35*, 363-379.

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Segment, gesture, prosody* (pp. 68-86). New York: Cambridge University Press.

Bernstein, L. E., Takayanagi, S., & Auer, E. T., Jr. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Communication, 44*, 5-18.

Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother–child speech. *Journal of Phonetics, 14*, 303-309.

Bernstein Ratner, N., & Luberoff, A. (1984). Cues to post-vocalic voicing in mother-child speech. *Journal of Phonetics, 12*, 285-289.

Biringen, Z. C. (1987). Infant attention to facial expressions and facial motion. *Journal of Genetic Psychology, 148,* 137-33.

Boersma, P., & Weenink, D. (2009). Pratt: Doing phonetics by computer (Version 5.1.12) [Computer program]. Retrieved September 1, 2009, from http://www.praat.org/

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': Modifications in mothers' infant-directed action. *Developmental Science, 5,* 72-83.

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science, 296,* 1435.

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a,i/ in English. *Journal of the Acoustical Society of America, 117,* 3867-3878.

Chong, S. C. F., Werker, J. F., Russell, J. A., & Carroll, J. M. (2003). Three facial expressions mothers direct to their infants. *Infant and Child Development, 12,* 211-232.

Cohn, J. F., & Elmore, M. (1988). Effect of contingent changes in mothers' affective expression on the organization of behavior in 3-monthold infants. *Infant Behavior and Development, 11,* 493-505.

de Boer, B. (2003). *Emergence of sound systems through self-organization.* Vrije Universitiet Brussel.

Dominey, P. F., & Dodane, C. (2004). Indeterminacy in language acquisition: The role of child directed speech and joint attention. *Journal of Neurolinguistics, 17,* 121-145.

Dromey, C., & Ramig, L. O. (1998). The effect of lung volume on selected phonatory and articulatory variables. *Journal of Speech, Language, and Hearing Research, 41*, 491-502.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). Articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America, 89,* 369-382.

Englund, K., & Behne, D. (2006). Changes in infant directed speech in the first six months. *Infant and Child Development, 15,* 139-160.

Erickson, D. (1998). Effects of contrastive emphasis on jaw opening. *Phonetica, 55,* 147-169.

Erber, N. P. (1974). Effects of angle, distance, and illumination on visual reception of speech by profoundly deaf children. *Journal of Speech and Hearing Research, 17*, 99-112.

Fairbanks, G. (1950). A physiological correlative of vowel intensity. *Speech Monograph, 22,* 390-395.

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

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*, 104-113.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., 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.

Fenson, L., Dale, P., Reznick, S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59* (Serial No. 242).

Fromkin, V. (1964). Lip positions in American English vowels. *Language and Speech, 7*, 215-225.

Goffman, L., Smith, A., Heisler, L., & Ho, M. (2008). The breadth of coarticulatory units in children and adults. *Journal of Speech and Hearing Research,* 51(6), 1424-1437.

Gogate, L., & Bahrick, L. E. (2001). Intersensory redundancy and seven-month-old infants' memory for arbitrary syllable-object relations. *Infancy, 2*, 219-231.

Grant, K. W., & Seitz, P. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *Journal of the Acoustical Society of America, 108,* 1197-1208.

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24*, 14-20.

Hall, M. D., Green, J., Moore, C. A., & Kuhl, P. K. (1999). Contribution of articulatory kinematics to visual perception of stop consonants. In P. K. Kuhl & L. Crum (Eds.), *Proceedings 2nd Convention of the European Acoustics Association: Forum Acusticum and 137th Meeting of the Acoustical Society of America (#4aSCb15).* Woodbury, NY: Acoustical Society of America.

Harrington, J., Fletcher, J., & Roberts, C. (1995). Coarticulation and the accented/unaccented distinction. *Journal of Phonetics 23,* 305-322.

Jusczyk, P., Hohne, E., & Mandel, D. (1995). Picking up regularities in the sound structure of the native language. In W. Strange (Ed.), *Speech perception and linguistic experience issues in cross-language research* (pp. 91-119). Timonium, MD: York Press.

Kaplan, P. S., Bachorowski, J.-A., Smoski, M. J., & Hudenko, W. J. (2002). Infants of depressed mothers, although competent learners, fail to learn in response to their own mother's infant-directed speech. *Psychological Science, 13*, 268-71.

Katz, G. S., Cohn, J. F., & Moore, C. A. (1996). A combination of vocal F_0 dynamic and summary features discriminates between three pragmatic categories of infant directed speech. *Child Development, 67*, 205-217.

Kemler Nelson, D., Hirsh-Pasek, K., Jusczyk, P. W., & Wright Cassidy, K. (1989). How prosodic cues in motherese might assist language learning. *Journal of Child Language, 16*, 53-68.

Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America, 72*, 353–365.

Kitamura, C., & Burnham, D. (1998). The infant's response to maternal vocal affect. In C. Rovee-Collier (Ed.), *Advances in infancy research* (pp. 221-236). Norwood, NJ: Ablex.

Klatt, D.H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59*, 1208-1221.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*, 684-686.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science, 218*, 1138-1141.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403-409). Netherlands: Kluwer Academic.

Lindblom, B. E. F., & Sundberg, J. E. F. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustical Society of America, 50*, 1166-1179.

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

Locke, J. L. (1993). *The child's path to spoken language.* Cambridge, MA: Harvard University Press.

Matlab [Computer software]. (2009). Natick, MA: The MathWorks.

MacLeod, A., & Summerfield, A. (1987). Quantifying the contribution of vision to speech perception by noise. *The British Journal of Audiology, 21*, 131-141.

McRoberts, G. W., & Best, C. T. (1997). Accommodation in mean f0 during mother-infant and father-infant vocal interactions: A longitudinal case study. *Journal of Child Language, 24*, 719-736.

Mefferd, A. S., & Green, J. R. (in press). Articulatory-To-Acoustic Relations in Response to Speaking Rate and Loudness Manipulations. *Journal of Speech and Hearing Research.*

Milenkovic, P. (2004). *TF32* [Computer software]. University of Wisconsin-Madison, Department of Electrical and Computer Engineering.

Montgomery, A. A., & Jackson, P. L. (1983). Physical characteristics of the lips underlying vowel lipreading performance. *Journal of Acoustic Society of America, 73,* 2134-2144.

Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of Acoustic Society of America, 96,* 40-55.

Munhall, K., Jones, J., Callan, D., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility: Head movement improves auditory speech perception. *Psychological Science, 15,* 133-137.

Murray, L., & Trevarthen, C. (1985). Emotional regulations of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 177-197). Norwood, NJ: Ablex.

Newborg, J. (2005). *Battelle Developmental Inventory* (2nd ed.). Itaska, IL: Riverside.

Patterson, R. L., & Werker, J. F. (2003). Two-month-old infants match phonetic information in lips and voice. *Developmental Science, 6,* 191-196.

Payton, K., Uchanski, R., & Braida, L. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *The Journal of the Acoustical Society of America, 95,* 1581-1592.

Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1993). Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: a pilot "motor equivalence" study. *Journal of the Acoustical Society of America, 93,* 2948-2961.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research, 28,* 96-103.

Rvachew, S., Slawinski, E. B., Williams, M., & Green, C. (1996). Formant frequencies of vowels produced by infants with and without early onset otitis media. *Canadian Acoustics, 24,* 19-28.

Robb, M. P., Chen, Y., & Gilbert, H. R. (1997). Developmental aspects of formant frequency and bandwidth in infants and toddlers. *Folia Phoniatr Logop, 49,* 88–95.

Rosenblum, L. D., Johnson, J. A., & Saldaña, H. M. (1996). Visual kinematic information for embellishing speech in noise. *Journal of Speech and Hearing Research, 39,* 1159-1170.

Schulman, R. (1989). Articulatory dynamics of loud and normal speech. *Journal of the Acoustical Society of America, 85,* 295-312.

Smiljanic, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *Journal of the Acoustical Society of America, 118,* 1677-1688.

Slater, A., & Kirby, R. (1998). Innate and learned perceptual abilities in the newborn infant. *Experimental Brain Research, 123,* 90-94.

Stern, D. N. (1985). Affect attunement. In J. D. Call, E. Galenson, & R. L. Tyson (Eds.), *Frontiers of infant psychiatry (Vol. 2).* New York: Basic.

Stern, D. N., Spieker, 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. (1955). Development of a quantitative description of vowel articulation. *The Journal of the Acoustical Society of America, 27*, 401-493.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of Acoustical Society of America, 26,* 212–215.

Summerfield, A. Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In R. Campbell & B. Dodd (Eds.), *Hearing by eye* (pp. 3-51). Hillsdale: NJ: Lawrence Erlbaum.

Summerfield, Q., MacLeod, A. M., McGrath, M., & Brooke, N. M. (1989). Lips, teeth, and the benefits of lipreading. In A. W. Young & H. D. Ellis (Eds.), *Handbook of research on face processing* (pp. 223-233). Amsterdam: North-Holland.

Summers, W. V. (1987). Effects of stress and final-consonant noticing on vowel production articulatory and acoustic analyses. *The Journal of the Acoustical Society of America, 82*, 847-863.

Swanson, L. A., Leonard, L. B., & Gandour, J. (1992). Vowel duration in mothers' speech to young children. *Journal of Speech and Hearing Research, 35,* 617-625.

Tasko, S. M., & McClean, M. D. (2004). Variations in articulatory movement with changes in speech task. *Journal of Speech Language and Hearing Research, 47*, 85-100.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*, 53-71.

Toda, S., & Fogel, A. (1993). Infant response to the still-face situation at 3 and 6 months. *Developmental Psychology, 29,* 532-538.

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science, 11*, 188-195.

Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech and Hearing Research, 38*, 1001-1013.

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-S-H? A comparison of foreigner- and infant-directed speech. *Speech Communication, 17*, 2-7.

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences, USA, 102*, 1181-1186.

Wohlert, A., & Hammen, V. (2000). Lip muscle activity related to speech rate and loudness. *Journal of Speech Language and Hearing Research, 43*, 1229-1239.

Acknowledgements

This work has been supported by the National Institute of Health, National Institute on Deafness and Other Communication Disorders, Grants R03DC004643 and R01DC006463. The authors would like to thank Lacey LaBarge, Cynthia Didion, Cara Ullman, Lindsey Fairchild, Paige Mueller, and Kelly Raber for their assistance with data collection and analysis.

Table 1

Mean Vertical and Horizontal Lip Aperture in mm as a Function of Speaking Condition and Target Word.

	Vertical aperture		Horizontal aperture	
Stimulus	ADS	IDS	ADS	IDS
Reading passage	38.99 (5.90)	43.90 (6.57)	69.78 (5.56)	70.53 (5.76)
/ae/	35.17 (5.28)	37.59 (5.84)	63.94 (5.21)	63.99 (5.21)
/a/	31.08 (5.84)	32.72 (5.07)	63.67 (5.49)	63.69 (5.60)
/i/	28.17 (3.89)	28.94 (3.91)	65.28 (4.89)	65.08 (5.11)
/u/	24.30 (4.60)	24.86 (4.59)	63.60 (5.18)	63.53 (5.11)

ADS = adult directed speech, IDS = infant directed speech

FIGURE CAPTIONS

Figure 1. Hypothetical lip shape changes made by mothers to maximize the visual contrast among vowels. Each vowel is represented by its hypothetical location in lip shape space as defined by its vertical (lip opening) and horizontal lip (spread) aperture.

Figure 2. Schematic of laboratory setup during infant (top panel) and adult directed (bottom panel) speaking conditions.

Figure 3. Example of marker placements used to record lip movements.

Figure 4. Example of lip kinematics. The top window shows the downsampled acoustic recording of a mother saying “bat.” Vertical aperture, as displayed the middle window, represents the 3D Euclidean distance (mm) between the upper and lower lip markers. Horizontal aperture, as displayed in the bottom panel, represents the 3D Euclidean distance (mm) between the markers located near the left and right corners of the lips. The maximum vertical and horizontal apertures were recorded for each target vowel during the infant and adult directed speech.

Figure 5. The average change in fundamental frequency (Hz) and duration (ms). Error bars represent standard error across participants’ mean change in fundamental frequency.

Figure 6. Mean differences observed in vertical aperture (mm) between IDS and ADS. Error bars indicate the standard error of the mean. Means are displayed for all 25 mothers and for the 10 mothers who exhibited the greatest exaggeration in lip opening.

Figure 7. Maximum vertical aperture of the upper and lower lips for the vowel /a/ during IDS expressed as percentage of the change in vertical aperture from the oral closing from the consonant to oral opening during the ADS condition. For ease of interpretation, mothers were ranked from lowest to highest.

Figure 8. Acoustic vowel space observed during IDS and ADS. Error bars represent the standard error across participants' means.

Figure 1

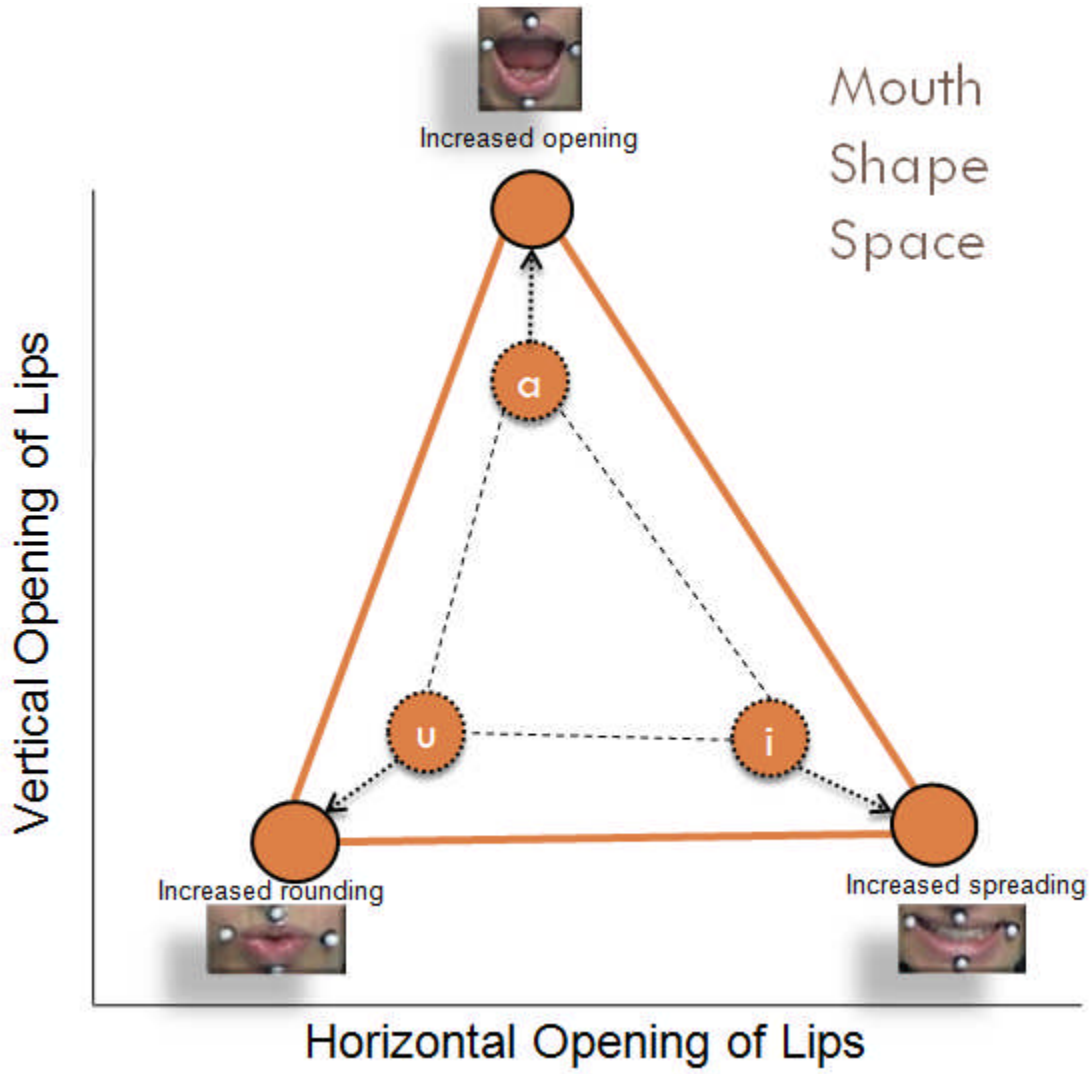

Mouth
Shape
Space
Increased opening
a
u
i
Increased rounding
Increased spreading
Vertical Opening of Lips
Horizontal Opening of Lips

Figure 2

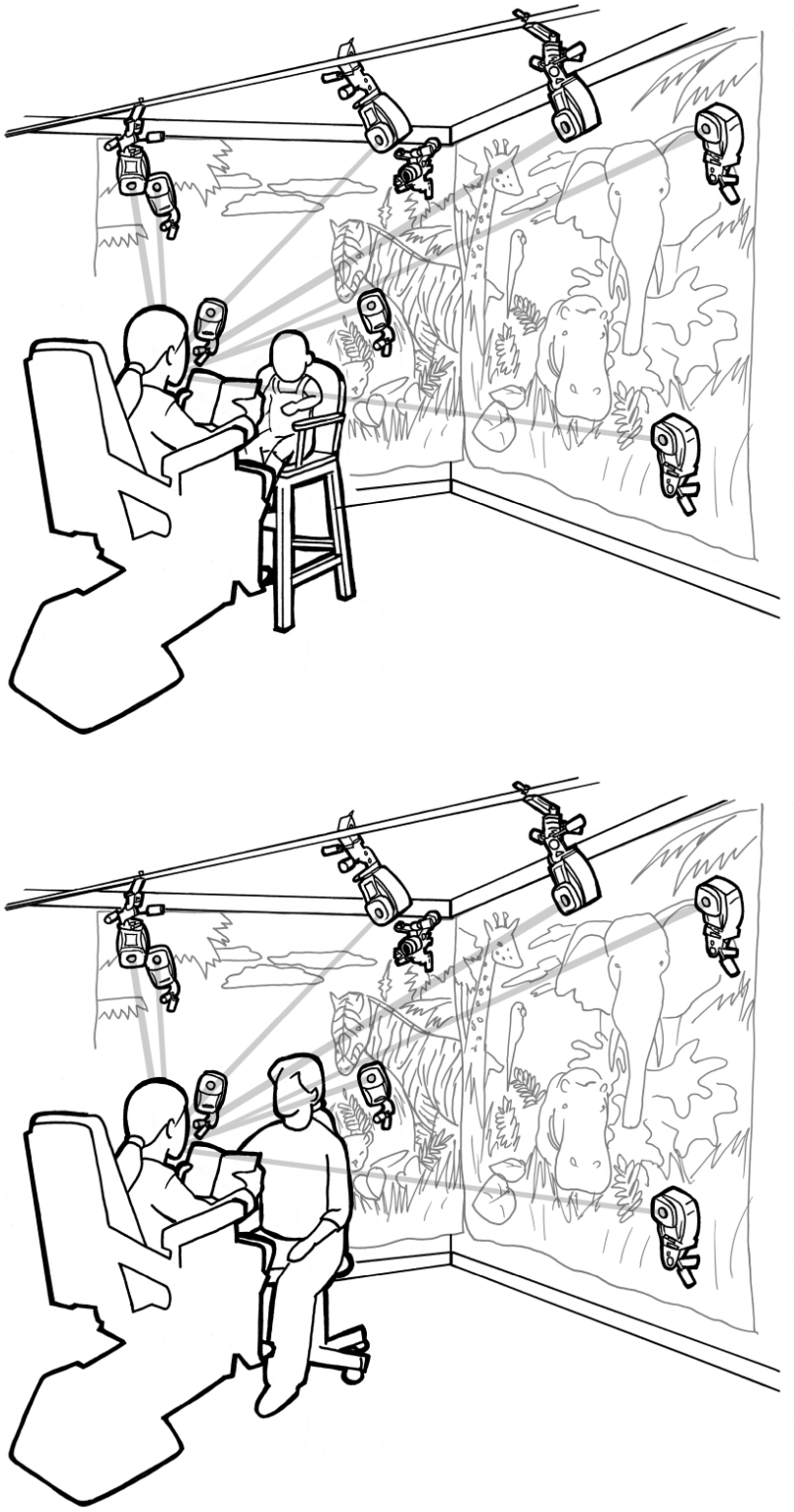
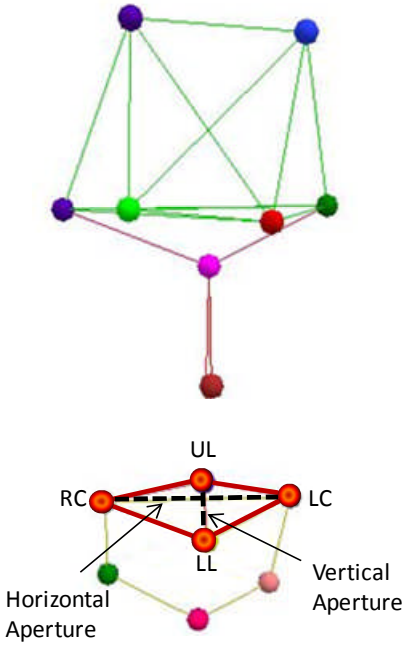


Figure 3



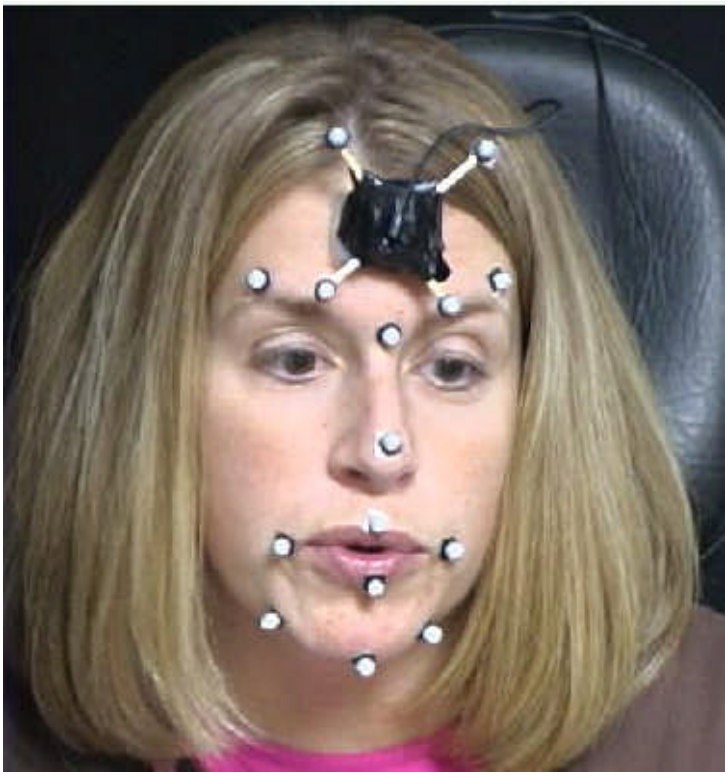


Figure 4

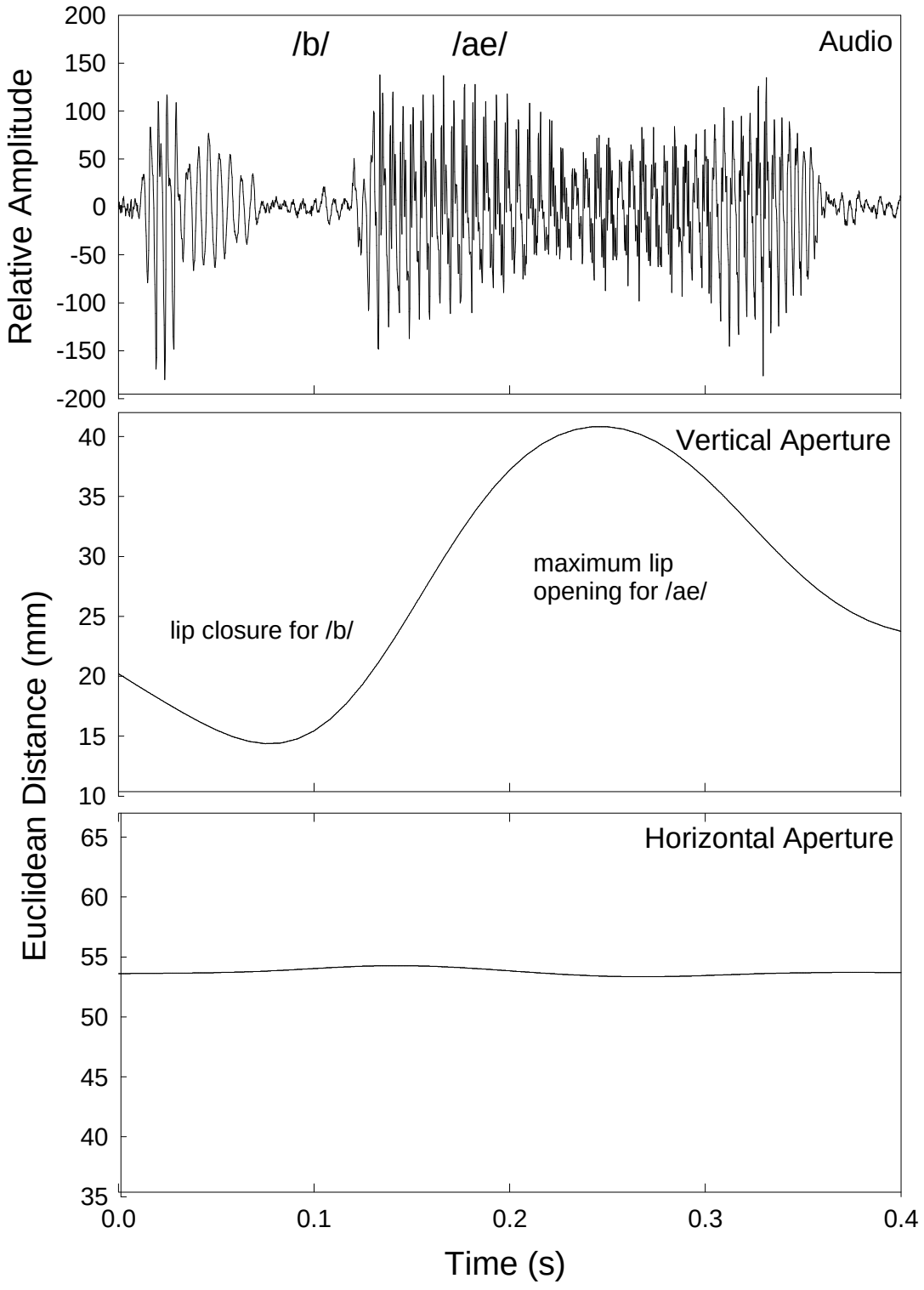

/b/
/ae/
Audio
Relative Amplitude
200
150
100
50
0
-50
-100
-150
-200
Vertical Aperture
maximum lip
opening for /ae/
lip closure for /b/
Euclidean Distance (mm)
40
35
30
25
20
15
10
Horizontal Aperture
65
60
55
50
45
40
35
0.0
0.1
0.2
0.3
0.4
Time (s)

Figure 5

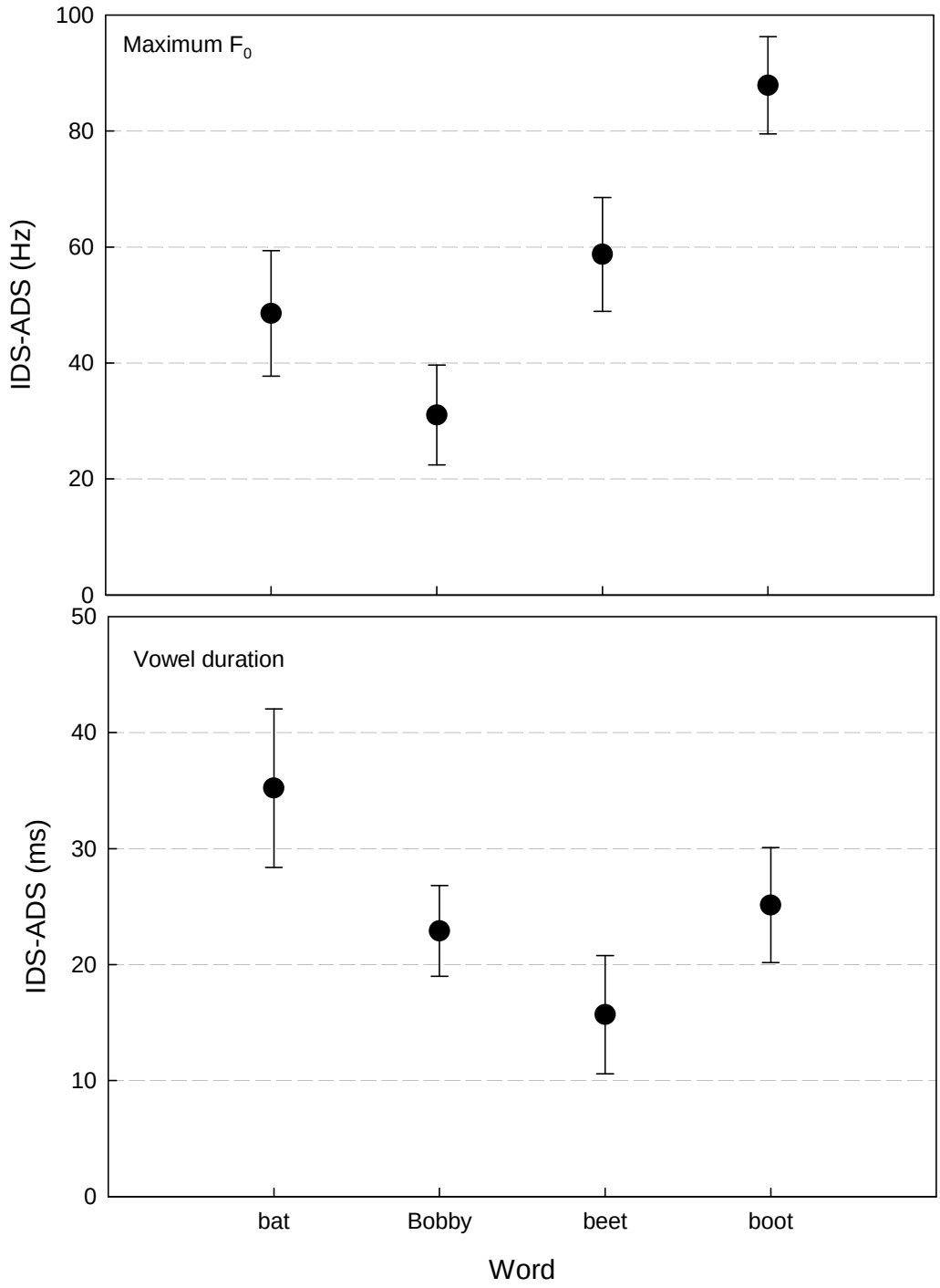

Maximum F_0
IDS-ADS (Hz)
100
80
60
40
20
0
Vowel duration
IDS-ADS (ms)
50
40
30
20
10
0
bat
Bobby
beet
boot
Word

Figure 6

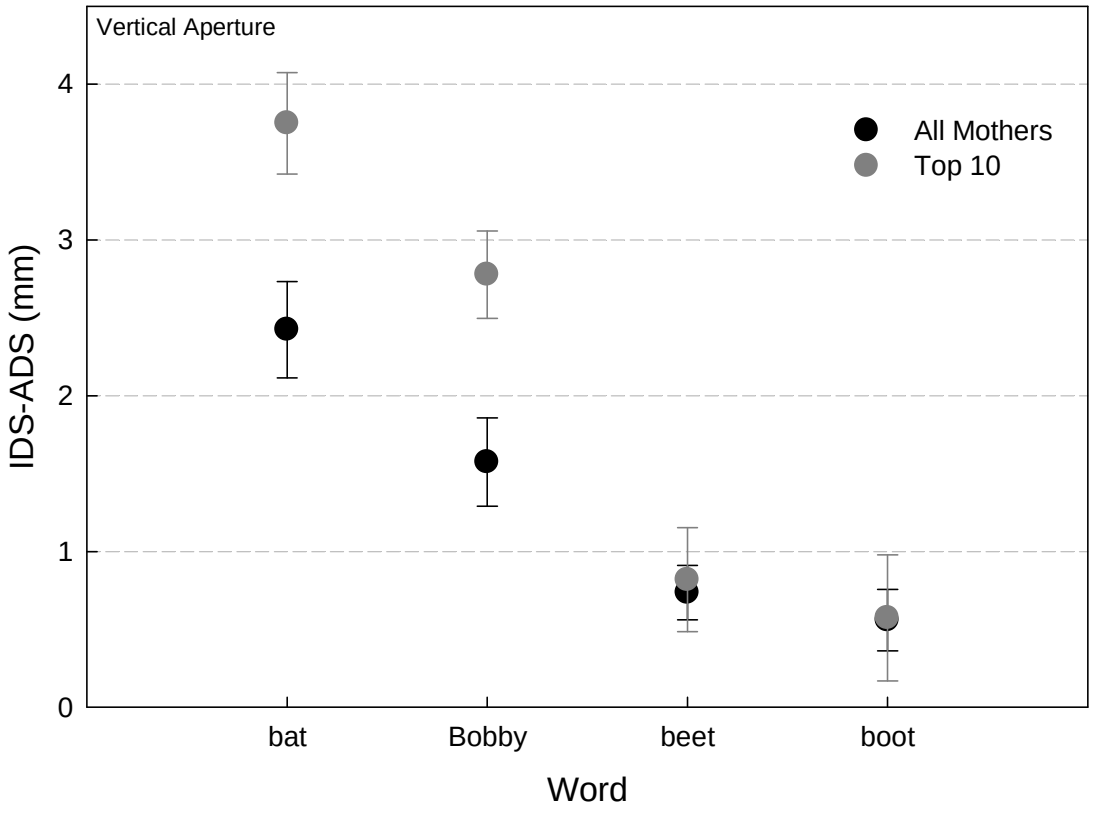

Vertical Aperture
IDS-ADS (mm)
0
1
2
3
4
All Mothers
Top 10
bat
Bobby
beet
boot
Word

Figure 7

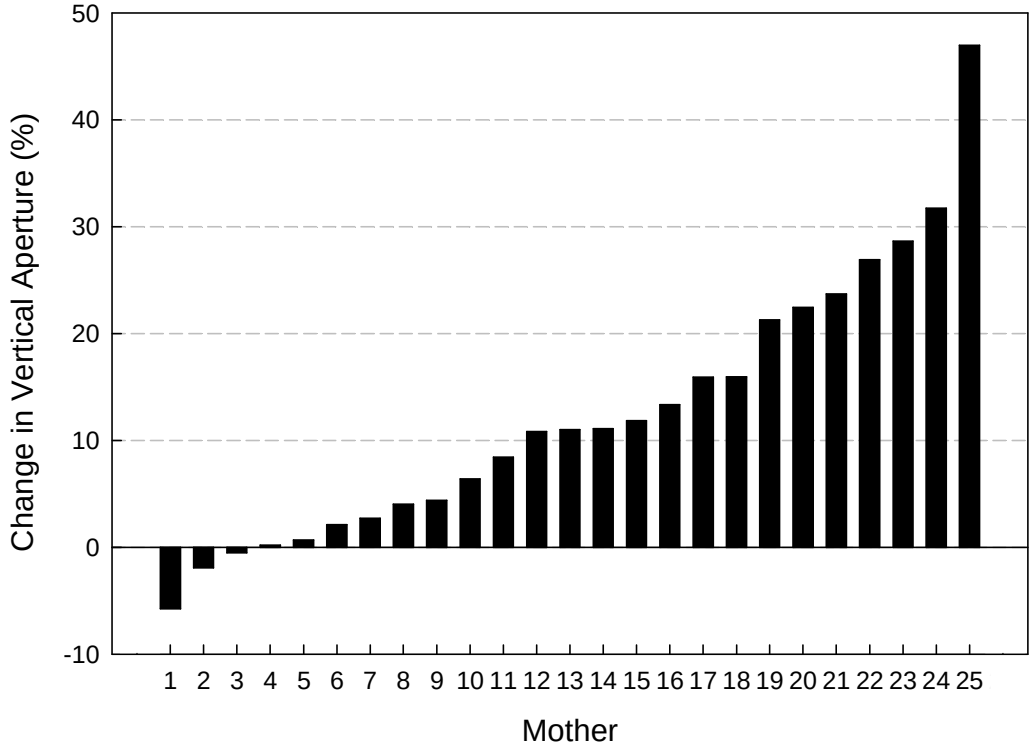
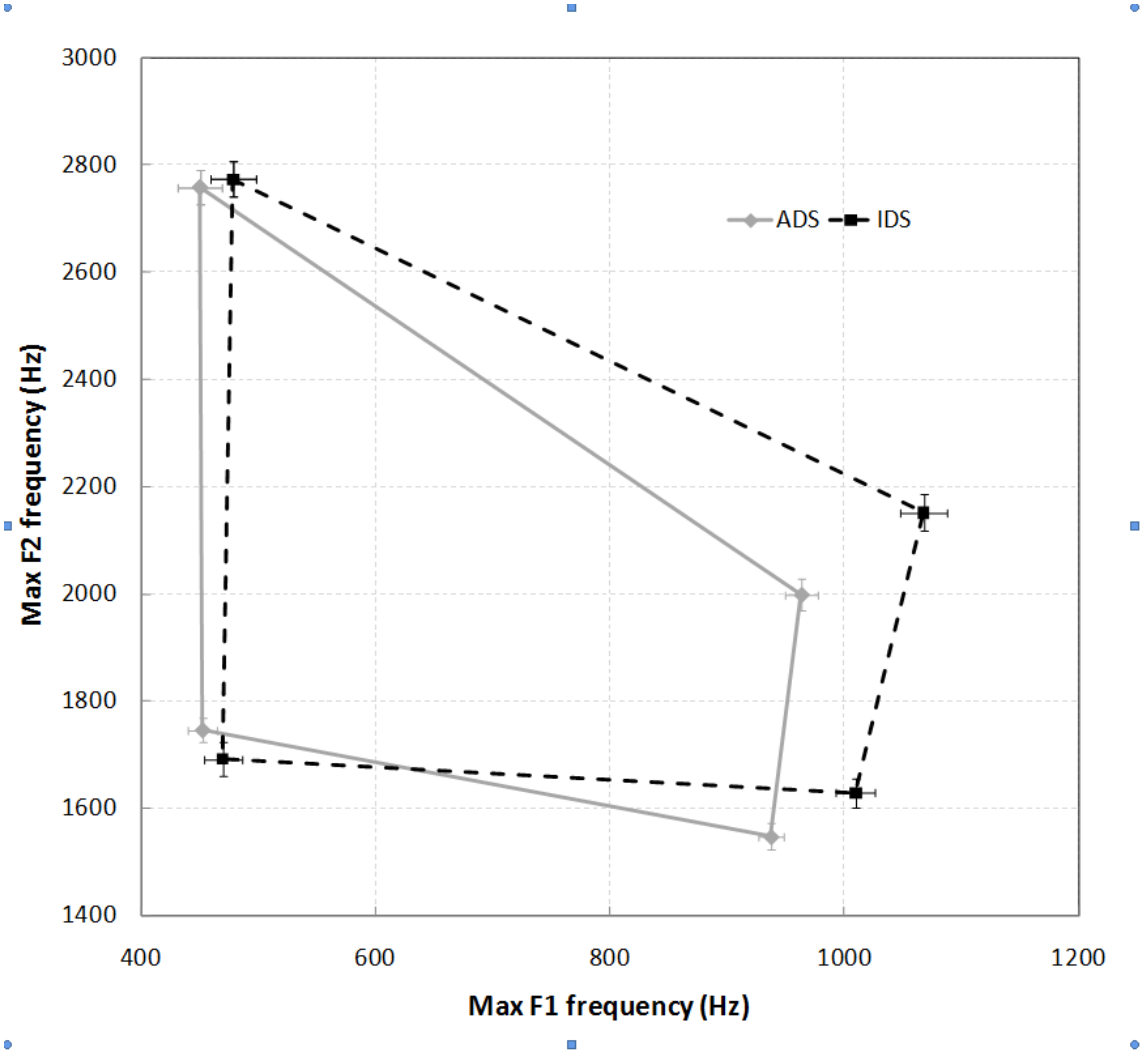

Change in Vertical Aperture (%)
50
40
30
20
10
0
-10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Mother

Figure 8


3000
2800
2600
2400
2200
2000
1800
1600
1400
400
600
800
1000
1200
Max F2 frequency (Hz)
Max F1 frequency (Hz)
ADS
IDS