Developmental Science

SHORT REPORT

Detecting 'infant-directedness' in face and voice

Hojin I. Kim and Scott P. Johnson

Department of Psychology, University of California, Los Angeles, USA

Abstract

Five- and 3-month-old infants' perception of infant-directed (ID) faces and the role of speech in perceiving faces were examined. Infants' eye movements were recorded as they viewed a series of two side-by-side talking faces, one infant-directed and one adult-directed (AD), while listening to ID speech, AD speech, or in silence. Infants showed consistently greater dwell time on ID faces vs. AD faces, and this ID face preference was consistent across all three sound conditions. ID speech resulted in higher looking overall, but it did not increase looking at the ID face per se. Together, these findings demonstrate that infants' preferences for ID speech extend to ID faces.

Introduction

The social environment of young infants is markedly different from that of adults. Adults often modify their communicative behaviors when interacting with infants, employing a speech register characterized by elevated fundamental frequency, wider intonation contours, more precise articulation, increased use of repetition, decreased complexity, elongated vowels, reduced speech rate, shorter phrases, and longer pauses (Fernald & Kuhl, 1987; Fernald & Simon, 1984; Kitamura & Burnham, 2003). Infants demonstrate robust preferences for infant-directed (ID) speech over adult-directed (AD) speech (Cooper & Aslin, 1990; Fernald 1985). ID speech has a special propensity to attract and hold infants' attention (Fernald & Simon, 1984), conveys affective intentions (Fernald, 1992), and facilitates language processing (e.g. Liu, Kuhl & Tsao, 2003; Thiessen, Hill & Saffran, 2005; Werker, Pons, Dietrich, Kajikawa, Fais & Amano, 2007).

Speech, including ID speech, is closely related to facial expressions. The muscles used to produce facial expressions influence the articulation of speech sounds (Massaro, 1998), and infants attend to certain visual cues, especially the mouthing shapes, in faces that produce speech. Kuhl and Meltzoff (1984) used a preferential looking paradigm to investigate 4-month-old infants' sensitivity to visual cues for the vowels /i/ and /a/. For both vowels, infants looked longer at the face that matched the vowel. Young infants imitate mouth movements when presented with compatible audiovisual representations of vowels (Legerstee, 1990), and are susceptible to the McGurk effect (Burnham & Dodd, 2004; Rosenblum, Schmuckler & Johnson, 1997), an auditory-visual illusion that illustrates how perceivers merge information for speech sounds across the senses (McGurk & MacDonald, 1976). Together these studies provide evidence that infants detect the congruency between facial movement and speech sounds.

Because face and voice present synchronous information, and because infants are sensitive to some of this information, it raises the question whether facial expressions, like speech, might be appropriately characterized as infant-directed when adults talk to infants. Stern (1974) reported that when speaking to their own infants, mothers' facial expressions were often more exaggerated, slower in tempo, and longer in duration than AD facial expressions. More recently, Chong, Werker, Russell and Carroll (2003) examined mothers' faces during mother–infant interactions and reported three distinct types of facial expression: soothing and comfort, amazement and pride, and exaggerated smiles.

In the present study, we asked whether infant preference for ID speech could extend to ID faces. Moreover, we asked whether infants are sensitive to the facial gestures as well as the auditory properties of ID vs. AD

Address for correspondence: Hojin Kim, 3261 S. Sepulveda Blvd, #103, Los Angeles, CA 90034, USA; e-mail: hojinkim@ucla.edu

speech. Young infants prefer ID over AD communication when exposed to talking faces (Werker, Pegg & McLeod, 1994) and action displays (Brand & Shallcross, 2008), but to our knowledge no previous study has (a) isolated faces and voices from the same segments of infant-directed communications, (b) investigated preference for ID faces alone, and then (c) added in the voices to examine modulation of these preferences. We recorded infants' eye movements as they viewed AD and ID faces side-by-side, in silence or accompanied by asynchronous AD or ID speech. We hypothesized that there would be a baseline preference for ID face over AD face (i.e. in silence), and, if infants recognize an ID register in both faces and voices, that infant attention toward the faces would be modulated according to the speech they heard during in-sound trials. This would be evinced as overcoming the baseline preference in favor of AD face when accompanied by AD speech, and even greater attention to ID face when accompanied by ID speech relative to baseline, the latter effects due to the tendency of infants to look longer at a visual-auditory match relative to a mismatch (cf. Bahrick, 1998). Thus our study examines infants' detection of 'infant-directedness' in face and voice when they are presented asynchronously.

In a longitudinal study of 30 infants, Lamb, Morrison and Malkin (1987) demonstrated that infants' face-to-face play with mothers becomes increasingly frequent and reaches a peak between 3 and 5 months of age. We reasoned, therefore, that 5-month-old infants may show a multimodal match based on register due to substantial exposure to infant-directed speech and faces, whereas 3-month-olds, who have more limited experience engaging in face-to-face interaction with adults, would not show the match.

Method

Participants

Forty-two full-term 5-month-olds (21 girls, 21 boys, $M = 5.17$ months, range = 4.2–5.9 months) and 33 full-term 3-month-olds (16 girls, 17 boys, M age = 3.11 months, age range = 2.5–3.6 months) participated in the study. Twenty-six infants were observed but excluded from the analysis due to fussiness (nine 5-month-olds, seven 3-month-olds) or equipment failure/experimenter error (10). Infants were recruited from birth records provided by the county. Parents were first sent a letter of invitation to participate in the experiment; interested parents returned a postcard and were later contacted by telephone. Parents were provided with a small gift for their infants but were not paid for participation.

Stimulus preparation

Infants viewed a pair of videotaped events accompanied by a vocal soundtrack. Each event showed a woman's face as she engaged in a live face-to-face interaction with a member of her own family – either her husband or her 18-month-old infant. One woman served as the model for all stimuli. The recording sessions yielded segments of video and audio records that were either adult- or infant-directed.

The model and family members viewed a video monitor showing the person with whom they were conversing during the live interaction (cf. Murray & Trevarthen, 1985). We asked the model to talk about identical topics (e.g. birthday parties, family vacation) to both listeners for approximately the same duration. Recording sessions of model and family members lasted between 5 and 10 minutes. The model's facial expressions and speech were recorded with a Sony digital camcorder subsequently connected to a Macintosh computer to save the recordings onto the hard drive.

Recordings were segmented into multiple 20-second clips with iMovie software and separated into video and audio files using Adobe Premiere. Sixteen undergraduates (half female and half male) rated the face and speech clips as infant- or adult-directed using a modified Likert scale, a value of 10 denoting a face or speech sample as 'definitely produced when she was interacting with an infant' and a value of 1 as 'definitely produced when she was interacting with an adult'. The six ID face clips and six AD face clips with the highest (most extreme ID) and lowest (most extreme AD) ratings, respectively, were selected to make six side-by-side face stimuli with Adobe Premiere. Overall both the ID speech and the ID face clips were judged to have been more likely produced when the model interacted with an infant, confirming the infant-directedness of our stimuli (see Figure 1). Left–right presentation of the ID face was counterbalanced across presentations within each of the sound conditions. Two ID and two AD speech clips were also chosen based on the adults' ratings. One speech clip was added to each side-by-side face stimulus for a total of eight (four face pairs with ID speech, two with the ID face on the left and two with the ID face on the right, and the same arrangement for the four face pairs with AD speech), and four silent side-by-side face stimuli were created (two with the ID face on the left and two with the ID face on the right).

Stimulus properties

The soundtrack was asynchronous with both faces during in-sound trials so as to rule out the possibility
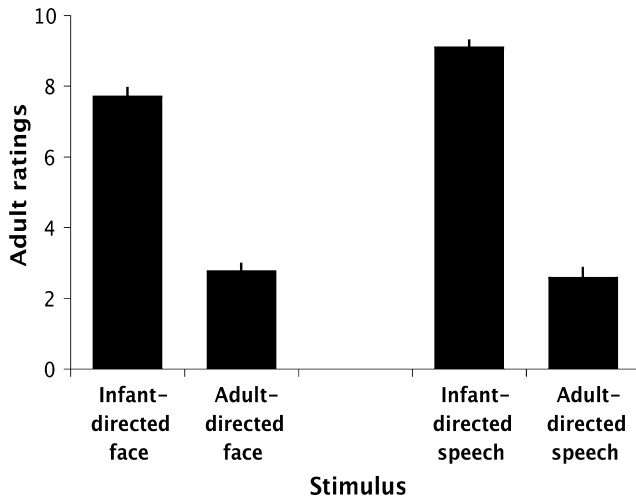
Figure 1 *Adult ratings of the six face and four voice stimuli. Faces recorded as the model interacted with her child were rated as more infant-directed, and faces recorded as she interacted with her husband were rated as more adult-directed (t(15) = 13.25, p < .0001). The same was true for recordings of the voice (t(16) = 16.59, p < .0001).*

that infants used temporal synchrony to match face and voice. Stimuli were presented in one of four pseudorandom orders with the stipulation that no more than two of any sound condition (ID speech, AD speech, no speech) could be presented in succession. Each visual stimulus measured 15.2 × 15.9 cm (14.4 × 15.1° visual angle) and was separated by a gap of 2.5 cm (2.4°). Each face measured approximately 10.2 × 7.6 cm (9.7 × 7.3°). See Figure 2 for an example.

We reasoned that infants' preferences might stem from differences in infants' interpretation of the emotional expressions in both ID and AD faces, assuming that ID faces in general are 'happier' (as ID speech may be

Figure 2 *Examples of infant-directed face (left) and adult-directed face (right). Ovals denote areas of interest (AOI) within which infant dwell times were recorded.*

'happier' than AD speech; Singh, Morgan & Best, 2002), even if, as in the present study, the model interacted with a close family member, exhibiting consistently positive emotional responses. To address the possibility of overlap in infant-directedness and happiness, we obtained adult ratings of 'happy' vs. 'neutral' in our face and voice stimuli in the same fashion as reported for ratings of infant- vs. adult-directedness (Figure 1). If infant-directedness is undifferentiated from happiness, these ratings should be statistically similar. However, ratings of infant-directedness were consistently higher than ratings of happiness in both ID faces, t(15) = 3.69, p < .01, and ID speech, t(15) = 6.30, p < .001, implying that infant-directedness and happiness are distinguishable, at least for adults.

The dependent variable in this experiment was the mean dwell time on two side-by-side dynamic face stimuli. On each trial, one of three speech types (ID, AD, and no speech) was presented with two (ID and AD) faces. Each infant was exposed to all three speech types and two faces, presented twice for a total of 12 stimuli, each 20 s in duration.

Procedures

Infants were tested individually, seated on a parent's lap 60 cm from a 17-inch monitor surrounded by black curtains. Eye movements were recorded with a Tobii ET-1750 eye tracker at 60 Hz with a spatial accuracy of approximately 1° (cf. Morgante, Zolfaghari & Johnson, 2012). The soundtrack came from speakers located behind the monitor. The lights in the experimental room were dimmed and the only source of illumination came from the monitor.

To calibrate each infant's point of gaze, a dynamic target-patterned ball undergoing repeated contraction and expansion around a central point was presented briefly at five locations on the screen (the four corners plus the center) as the infant watched. The Tobii eye tracker provides information about calibration quality for each point; if there were no data for one or more points or if calibration quality was poor, calibration at those points was repeated. Calibration was followed immediately by presentation of faces as described previously. Prior to each trial a small attention-getting stimulus was shown to re-center the point of gaze.

Results

The data consisted of the mean dwell time (DT) per infant on each face. We report here only DT within the faces, recorded by superimposing areas of interest on the

faces using Clearview software (see Figure 2). Areas of interest were generously sized, extending approximately 3° beyond the face, to accommodate head movement of the model as she spoke. Infants on average accumulated 135.3 s of total DT on the faces across the 12 trials (range = 60.4–213.7 s). Thirty-nine of the 5-month-olds included in the analyses completed all 12 trials; two completed 11 trials, and one completed six trials (a complete set of each trial type). Twenty-six of the 3-month-olds completed all 12 trials and seven completed six trials. Preliminary analysis of possible effects of Sex on looking behaviors revealed no significant main effects or interactions that bore on the principal questions of interest; therefore data were collapsed across this variable in the primary analyses.[1]

Our principal research questions concerned (a) whether infants might have a visual preference for ID faces, and (b) whether such preferences might develop between 3 and 5 months. We also asked whether infants may be sensitive to the amodal nature of infant-directedness, in which case an ID face preference might vary as a function of (asynchronous) ID or AD speech. Therefore we considered the silent condition a 'baseline' starting point to examine these possibilities, and against which to evaluate modulation of preferences by asynchronous ID or AD speech. As seen in Figures 3 and 4, infants at both ages tended to look longer at ID faces regardless of sound condition. However, t-tests revealed that DTs were reliably greater in ID faces only in the 5-month-old group in silent and ID speech conditions, $t(41) = 5.05$, $p < .0001$, and $t(41) = 4.46$, $p < .001$, respectively (p-values Bonferroni corrected); the four remaining analyses were not statistically reliable when subjected to Bonferroni corrections for multiple comparisons.

To provide direct comparisons of performance across conditions and age, we computed a 2 (Age Group) × 3 (Sound: no speech (silence), ID speech, or AD speech) × 2 (Face: ID vs. AD) mixed ANOVA, with repeated measures on the second and third factors. This analysis revealed a statistically significant main effect of Sound, $F(2, 146) = 4.83$, $p < .01$, partial $\eta^2 = .06$, the result of differences in accumulated DT when infants heard ID speech ($M = 23.4$ s, $SD = 7.6$), AD speech ($M = 22.2$ s,
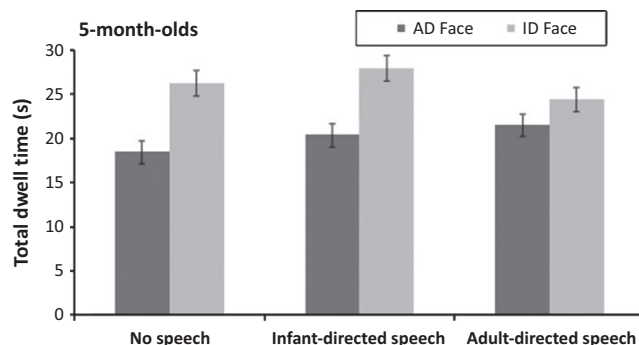

Figure 3 *Data from 5-month-olds showing the mean dwell time on each type of face stimulus (ID and AD) under the three speech conditions (no speech, ID speech, and AD speech). Error bars = SEM.*
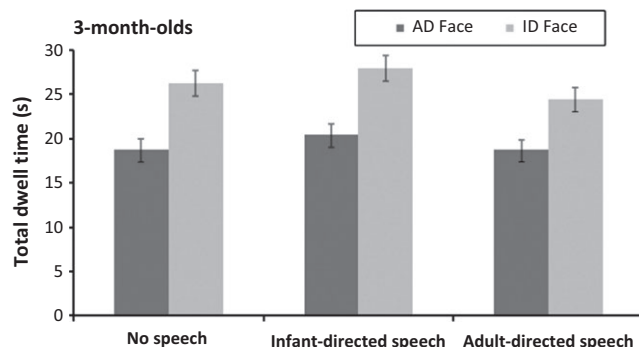

Figure 4 *Data from 3-month-olds showing the mean dwell time on each type of face stimulus (ID and AD) under the three speech conditions (no speech, ID speech, and AD speech). Error bars = SEM.*

$SD = 6.5$), or in silence ($M = 21.8$ s, $SD = 7.1$). Tests of simple effects revealed that DTs in the ID speech condition were greater than those in the silent condition, $F(1, 73) = 10.24$, $p < .01$, and in the AD speech condition, $F(1, 73) = 4.62$, $p < .05$. There was also a reliable main effect of Face, $F(1, 73) = 48.58$, $p < .00001$, partial $\eta^2 = .40$, due to overall higher DTs in ID faces relative to AD faces. There were no other significant main effects or interactions.

We also computed a series of planned comparisons (simple effects tests) targeted toward questions of modulation of preferences according to sound within each age group (i.e. comparisons to baseline). For 3-month-olds, comparisons of performance in silence vs. AD and ID speech conditions yielded no reliable effects, ps > .80. For 5-month-olds, the silence vs. AD speech comparison yielded a marginally significant Sound × Face interaction $F(1, 73) = 3.84$, $p = .054$; and the comparison of silence vs. ID speech was not reliable, $p > .94$. A final set

---

[1] Preliminary analysis (a Sex × Age Group × Sound × Face mixed ANOVA) yielded an interaction between Sex and Age Group, $F(1, 71) = 5.36$, $p < .05$, partial $\eta^2 = .07$. This interaction appears to have resulted from a nonsignificant tendency for 3-month-old boys to look longer overall than 3-month-old girls; in 5-month-olds this tendency was reversed, but again was not statistically significant (the reasons for these effects are unclear). There were no reliable effects of Sex involving dwell times in ID vs. AD faces, or as a function of Sound.

of planned comparisons examined age differences in performance across the three sound conditions; none revealed a reliable effect, ps > .25.

Finally, we reasoned that ID faces, like ID speech, might be characterized by properties such as repetition, slower pace, and decreased complexity (relative to AD faces and voices), and that these properties might influence infants' preferences. Word counts for each of the 12 face stimuli in our experiment (six ID faces, six AD faces), revealed that the ID faces were slower ($M = 42.17$, $SD = 7.96$) than the AD faces ($M = 55.33$, $SD = 7.09$), $t(22) = 4.28$, $p < .0001$. To examine the possibility that the slower pace of ID faces was responsible for infants' preference for ID faces over AD faces that we report subsequently, three trials with similar word counts were compared: The difference in the number of words in these ID and AD faces was less than .33. Fifty-five of the 75 infants in the study produced DT on both faces in all three trials, and their DT was examined in a separate analysis. A 2 (Face: ID or AD) × 3 (Speech: no speech, ID speech, AD speech) within-subjects ANOVA revealed a significant main effect of Face, $F(1, 54) = 29.34$, $p < .001$, partial $\eta^2 = .35$, indicating reliably greater DT on the ID face ($M = 7.07$, $SD = 2.79$) than on the AD face ($M = 4.52$, $SD = 2.02$), and no other statistically significant effects. This analysis confirms that preferences for ID faces are not likely due simply to the decreased number of words that may typify ID communications.

In summary, we observed a preference for ID faces in 3- and 5-month-olds that did not appear to be age-dependent. When listening to ID speech, infants at both ages looked longer overall (i.e. at both faces) relative to AD speech and silent conditions. There is little evidence, however, that the ID face preference was facilitated by ID speech; evidence for attenuation of the ID face preference when listening to AD speech was inconclusive in 5-month-olds, and absent altogether in 3-month-olds.

Discussion

Infants at 3 and 5 months looked reliably longer at a model's face recorded when she interacted with her own infant, relative to interactions with her husband, as the faces were shown side-by-side. We interpreted this finding as evidence that young infants prefer infant-directed faces relative to adult-directed faces, perhaps analogous to young infants' preference for ID speech relative to AD speech. We obtained little evidence that the ID face preference was modulated by the variations in speech we employed in our study – whether silent, AD speech, or ID speech – although ID speech had the effect

of increasing looking overall. (The use of asynchronous audiovisual stimuli might have contributed to the absence of any facilitating effect; see Kahana-Kalman & Walker-Andrews, 2001).

These results are consistent with the first of our two predictions. First, we hypothesized that infants would show greater visual attention toward ID faces than AD faces. This hypothesis was supported both by overall higher dwell time on ID faces and by baseline preferences for ID faces in silence. Second, we predicted that the baseline preference would be modulated by the presence of speech that is consistent with one of the two faces – either AD or ID – and that this would be revealed by greater dwell time on the face congruent with voice quality. When infants heard AD or ID speech, however, attention toward the congruent face was not consistently and reliably different from baseline. It may be that the ID face and voice were not sufficiently exaggerated relative to AD face and voice to merit increased attention on the infants' part, a possibility consistent with documented changes in the quality of IDS as infants age (e.g. Stern, Spieker, Barnett & MacKain, 1983). However, adults' ratings for ID face and speech provide evidence that our stimuli were effectively infant-directed, and the infants' clear preference for ID faces confirms this suggestion. We believe instead that the most likely explanation for comparable performance in baseline vs. ID speech conditions is that preferences in both instances were at ceiling, evidence for a strong inclination to attend to ID faces whether in silence or accompanied by a compatible soundtrack. It may be as well that AD faces are not sufficiently compelling to attract infants' attention when ID faces are available, even in the presence of a matching soundtrack; alternatively, it may be that intermodal matching under tested conditions is not possible in infants so young.

By 3 months, therefore, infants detect and process visual features present in ID faces. To our knowledge, the present results are the first to document infants' preference for ID faces, and as such these findings support and extend previous reports of adults' ID behaviors, many of which are multimodal, involving vocal, gestural, and facial expressions presented in synchrony (Gogate, Bahrick & Watson, 2000). Other studies have identified ID sign language (Masataka, 1992) and ID action (Brand, Baldwin & Ashburn, 2002); these also elicit preference over their AD counterparts (Masataka, 1996; Brand & Shallcross, 2008).

It is unlikely that extensive experience – exposure to infant-directedness – is the sole means by which infants come to prefer ID speech, given that these preferences have been observed at birth (e.g. Cooper & Aslin, 1990). Whether experience viewing ID faces leads to a preference by 3 months is not known, but such a possibility is

consistent with adults' engagement of infant attention via ID behaviors from an early age (Lamb *et al.*, 1987). Alternatively, it is possible that specific features of ID faces, distinct from the defining features of ID speech that might also be present in facial movements (such as repetition, slower pace, and decreased complexity), were responsible for infants' preference. However, our data are clear in their indication that preference for ID faces is not solely determined by pacing, because pacing across ID and AD faces in these trials was approximately equal.

ID faces are characterized by wider smiles and eye constriction stemming from raised cheeks, presumably resulting from heightened emotional content (Messinger, Mahoor, Chow & Cohn, 2009). ID faces also exhibit greater eye contact (Brand, Shallcross, Sabatos & Massie, 2009), and mutual gaze is an attractive stimulus from birth (Farroni, Csibra, Simion & Johnson, 2002). For our stimuli, the model was instructed to maintain gaze on her partner during the recordings, eliminating differences in eye contact as a possible basis for ID face preference, but the role of more subtle differences between the ID and AD faces in attracting infants' attention remains unknown. Of particular interest would be infants' interpretation of the emotional expressions in both ID and AD faces, assuming that ID faces in general are 'happier' (as ID speech may be 'happier' than AD speech; Singh *et al.*, 2002), even if, as in the present study, the model interacted with a close family member, exhibiting consistently positive emotional responses. Recently, we examined 6-month-old infants' visual preference for silent ID vs. AD faces and introduced controls for positive emotion (Kim & Johnson, 2013). The infants showed no preference for ID faces when both conveyed happiness, but a second group of infants looked significantly longer at AD faces conveying happy emotion over sad ID faces. These findings imply that infants' visual preference for ID faces is mediated, at least in part, by the presence of happy emotion.

ID behaviors may serve an important role as infants learn to discriminate between different emotional states in others. Kaplan, Jung, Ryther and Zarlengo-Strouse (1996) found that 4-month-olds exhibited increased visual attention for a neutral stimulus following a pairing of ID speech with a static happy face; AD speech had little effect, implying that the infants learned to associate ID speech with positive facial expressions. These effects are consistent with our finding of increased looking overall in the presence of ID speech. Notably, however, looking toward the mouth region when viewing talking faces increases later in infancy than the infants we tested (Lewkowicz & Hansen-Tift, 2012). This may allow infants to gain access to synchronous audiovisual speech information in support of language acquisition.

Four-month-olds also learned associations between 'consoling' ID speech and a static sad face, but not a happy face, suggesting that they formed selective associations between distinct emotions conveyed in speech and face (Kaplan, Zarlengo-Strouse, Kirk & Angel, 1997). ID behaviors, through their arousing effects in infants, could serve a functional role in assisting infants to respond to referential communication directed to them (Senju & Csibra, 2008). Through these interactions, infants become increasingly sensitive to the context-specific nature of speech, facial expression, and other social behaviors, including both ID and AD behaviors, and perhaps come to better understand their own role as social participants.

Acknowledgements

References

Bahrick, L.E. (1998). Intermodal perception of adult and child faces and voices by infants. *Child Development*, **69**, 1263–1275.

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

Brand, R.J., & Shallcross, W.L. (2008). Infants prefer motionese to adult-directed action. *Developmental Science*, **11**, 853–861.

Brand, R.J., Shallcross, W.L., Sabatos, M.G., & Massie, K.P. (2009). Fine-grained analysis of motionese: eye gaze, object exchanges, and action units in infant- versus adult-directed action. *Infancy*, **11**, 203–214.

Burnham, D., & Dodd, B. (2004). Auditory-visual speech integration by pre-linguistic infants: perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology*, **44**, 204–220.

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.

Cooper, R.P., & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, **61**, 1584–1595.

Farroni, T., Csibra, G., Simion, F., & Johnson, M.H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, USA*, **99**, 9602–9605.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior & Development*, **8**, 181–195.

Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals. In J. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). Oxford: Oxford University Press.

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

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

Gogate, L.J., Bahrick, L.E., & Watson, J.D. (2000). A study of multimodal motherese: the role of temporal synchrony between verbal labels and gestures. *Child Development*, **71**, 878–894.

Kahana-Kalman, R., & Walker-Andrews, A. (2001). The role of person familiarity in young infants' perception of emotional expressions. *Child Development*, **72**, 352–369.

Kaplan, P.S., Jung, P.C., Ryther, J.S., & Zarlengo-Strouse, P. (1996). Infant-directed versus adult-directed speech as signals for faces. *Developmental Psychology*, **32**, 880–891.

Kaplan, P.S., Zarlengo-Strouse, P., Kirk, L.S., & Angel, C.L. (1997). Selective and nonselective associations between speech segments and faces in human infants. *Developmental Psychology*, **33**, 990–999.

Kim, H., & Johnson, S.P. (2013). Do young infants prefer an infant-directed face or a happy face? *International Journal of Behavioral Development*, **37**, 125–130.

Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mothers' speech: adjustments for age and sex in the first year. *Infancy*, **4**, 85–110.

Kuhl, P.K., & Meltzoff, A.N. (1984). The intermodal representation of speech in infants. *Infant Behavior & Development*, **7**, 361–381.

Lamb, M.E., Morrison, D.C., & Malkin, M. (1987). The development of infant social expectations in face-to-face interaction: a longitudinal study. *Merrill-Palmer Quarterly*, **33**, 241–254.

Legerstee, M. (1990). Infants use multimodal information to imitate speech sounds. *Infant Behavior & Development*, **13**, 343–354.

Lewkowicz, D.J., & Hansen-Tift, A.M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences, USA*, **109**, 1431–1436.

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

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, **264**, 746–748.

Masataka, N. (1992). Motherese in signed language. *Infant Behavior & Development*, **15**, 453–460.

Masataka, N. (1996). Perception of motherese in a sign language by 6-month-old deaf infants. *Developmental Psychology*, **32**, 874–879.

Massaro, D. (1998). *Perceiving talking faces: From speech perception to a behavioral principle.* Cambridge, MA: MIT Press.

Messinger, D.S., Mahoor, M.H., Chow, S., & Cohn, J.F. (2009). Automated measurement of facial expression in infant–mother interaction: a pilot study. *Infancy*, **14**, 285–305.

Morgante, J.D., Zolfaghari, R., & Johnson, S.P. (2012). A critical test of temporal and spatial accuracy of the Tobii T60XL eye tracker. *Infancy*, **17**, 9–32.

Murray, L., & Trevarthen, C. (1985). Emotional regulation 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 Publishers.

Rosenblum, L.D., Schmuckler, M.A., & Johnson, J.A. (1997). The McGurk effect in infants. *Perception & Psychophysics*, **59**, 347–357.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, **18**, 668–671.

Singh, L., Morgan, J.L., & Best, C.T. (2002). Infants' listening preferences: baby talk or happy talk? *Infancy*, **3**, 365–394.

Stern, D.N. (1974). Mother and infant at play: the dyadic interaction involving facial, vocal, and gaze behaviors. In M. Lewis & L. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 187–232). New York: Wiley.

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.

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

Werker, J.F., Pegg, J.E., & McLeod, P.J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior & Development*, **17**, 323–333.

Werker, J.F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, **103**, 147–162.