



Categorization of Infant-Directed Speech: Processing Asynchronous Audio-Visual Speech



Kristin Kuhlman Atchison, Kate Georgelas Shepard and Melanie J. Spence
The University of Texas at Dallas - School of Behavioral and Brain Sciences

Introduction

Infant-directed speech (IDS) is a specialized speech register used by adults and children when speaking to infants. Typically, IDS has higher mean fundamental frequency, broader F₀ range, repetition of phrases and prosodic patterns, hyper-articulated vowels and simpler syntax than adult-directed speech (ADS) (Ferguson, 1964; Fernald, 1992; Fernald & Simon, 1994; Kuhl & Andruski, 1997).

The acoustic properties of IDS vary as a function of pragmatic context, our intent, and affect (Fernald, 1992). Different prosodic patterns have been identified as occurring in specific contexts, such as approving, comforting, prohibiting, and attention-getting. If an infant is upset, adults tend to speak in comforting utterances, whereas when adults seek to praise or encourage behavior an infant might hear approving utterances. IDS patterns can also be distinguished by dynamic and summary acoustic features. Dynamic acoustic features such as frequency contours, and summary features such as mean fundamental frequency, or frequency range within utterances, differentiate approvals and comforts (Katz, Cohen & Moore, 1996). Approving tones are typically higher in fundamental frequency than comforts and have more dynamic frequency contours.

The differences in how we speak to infants, either as compared to adults or as a function of our intent, may serve specific developmental functions for infants. Anne Fernald (1992) discusses three functions of IDS for infants at different points in development: 1) to modulate infant attention and affect, 2) to communicate caregiver intent and meaning, and 3) to exaggerate phonetic and linguistic structure of language, promoting learning.

Categorization of IDS is highly relevant to Fernald's second function of ID speech. Category formation, in which exemplars within a category are more likely to be treated as functionally equivalent than exemplars from another category, allows for reduced cognitive loads (Roch, Mervis, Gray, Johnson & Boyes-Braem, 1976). For IDS to communicate meaning, infants must first be able to distinguish between the categories of intent. For IDS to communicate meaning, infants must categorize ID utterances that convey different speaker intent; they must detect similarities between utterances conveying similar intent while disregarding irrelevant differences such as individual talker, volume, and rate of speech. If pre-verbal infants categorize IDS, they may begin to extract meaning from speech without comprehension of words.

Background and Aims

Previous research investigating IDS categorization by preverbal infants has produced differential results as a function of visual stimuli. When presented a checkerboard image while listening to IDS, 6-month-olds categorized approving and comforting IDS (Spence and Moore, 2003) but failed to categorize when shown a static face or a moving, talking and smiling face of an ADS video (Atchison, Spence & Touchstone, 2008). Infants also categorize IDS when they viewed the natural audio-visual synchronous video format (Atchison, Spence & Touchstone, 2009).

These findings suggest 6-month-old infants can categorize IDS with just the auditory information or the auditory and facial information together as presented in natural video format. Yet, when presented with a static or asynchronous adult-directed face, their categorization abilities are interrupted. It may be that 6-month-old infants' attention is diverted to the asynchrony between the facial and vocal speech, which interferes with their attention to the intent category boundaries of the speech samples.

To further explore the effect of facial-vocal asynchrony on IDS categorization, two types of asynchrony were examined. Experiment 1 examined IDS categorization when different communicative intents were conveyed by facial and vocal information (e.g., comforting facial expressions were paired with approving vocalizations), so there was a mismatch in intent and temporal synchrony. Experiment 2 examined IDS categorization when the same communicative intent was conveyed by facial and vocal information (e.g., both approving) but the audio and video streams were not synchronized.

Experiment 1:

Audiovisual Temporal Asynchrony & IDS Intent Asynchrony

Methods:

Participants: 6-month-olds: $N = 20$; $M = 181.15$ days old; Range = 165 – 192 days

Stimuli:

40 videos taken from 10 female talkers, each producing 2 comforting and 2 approving utterances. Asynchronous pairs were created by pairing similar-length approving audio tracks with comforting video tracks for each talker, and vice versa.

Approvals: Mean $F_0 = 311.72$ (SD = 32.92) Comforts: Mean $F_0 = 265.05$ (SD = 40.80)

Procedure:

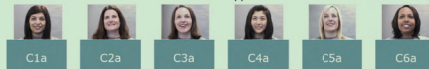
Each infant sat on parent's lap approximately 1 meter from 60" Sony HDTV



Infant-Controlled Procedure:
Control Event



Habituated to 6 videos of either approvals or comforts

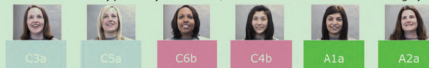


Six test trials: 2 post-habituation, 2 within-category, and 2 between-category test trials

Post-Habituation Trials: Repeated stimuli from habituation

Within-Category Test Trials: New exemplars from habituation category, spoken by previously seen talkers

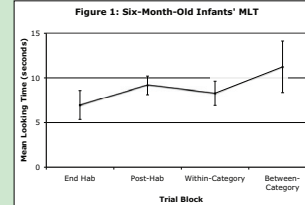
Between-Category Test Trials: New exemplars from novel category, spoken by previously seen talkers (order counterbalanced with within-category test trials)



Control Event



Results: No support for categorization was found.



No Main Effect of Trial Block: $F(3, 54) = 1.348$, $MSe = 47.601$, $p = 0.268$
Pairwise comparison (Within v. Between): Not significant

Experiment 2:

Audiovisual Temporal Asynchrony & IDS Intent Synchrony

Methods:

Participants:

6-month-olds: $N = 13$; $M = 181.69$ days; Range = 170 – 194 days

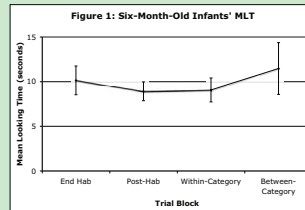
Stimuli:

Same as Experiment 1, but asynchronous pairs were matched within intent category as well as within talker. For example, the audio of stimuli "Good boy!" will be matched with the video of stimuli "Great job!" Thus stimuli were only mismatched on utterance and not intent category.

Procedure:

Identical to Experiment 1

Results: No support for categorization was found.



No Main Effect of Trial Block: $F(3, 33) = 0.488$, $MSe = 39.667$, $p = 0.693$
Pairwise comparison (Within v. Between): Not Significant

Discussion

When viewing either type of asynchronous stimuli, 6-month-olds failed to categorize approving and comforting IDS. Infants in both experiments did not look significantly longer to new exemplars from the novel category than new exemplars from the habituation category. No main effects of trial block or condition were seen. These results suggest 6-month-olds do not categorize IDS when the facial speech information conflicts with the audio speech in either temporal properties or communicative intent.

Conclusions

This research investigated if mismatched temporal properties or intent between facial and vocal IDS disrupt categorization of approving and comforting IDS. Six-month-olds were unable to categorize IDS when the intent and temporal properties of the facial speech did not match the audio speech in intent, but that was mismatched in temporal synchrony, did not support categorization (Experiment 2).

These results are consistent with previous research which has shown that infants' categorization of ID speech is affected by the type of visual stimulus infants see while hearing ID speech. Six-month-olds did not categorize approving and comforting ID speech when viewing a female static face and a silent asynchronous video of a face (Atchison, Spence & Touchstone, 2008). They do categorize when hearing ID speech without a facial stimulus (Spence & Moore, 2003) and when viewing audio-visual synchronous ID speech (Atchison, Spence & Touchstone, 2009). Asynchronous facial and vocal stimuli may interrupt 6-month-olds' attention to the category boundaries of IDS.

One function of IDS is communication of caregiver affect; adults' use of different IDS patterns in differing contexts may provide opportunities for infants to learn patterns of communicative intent. Development of IDS categorization may be prerequisite for this communicative function of IDS. This research illuminates the impact of facial speech on IDS categorization and is consistent with findings that 6-month-olds' IDS categorization is disrupted by asynchrony between facial and audio speech.

References

Atchison, K.K. & Spence, M.J. (2007, March). Four-month-old infants' categorization of infant-directed speech when viewing female, male, and scrambled faces. Poster session presented at the biennial meeting for the Society for Research in Child Development, Boston, MA.
Atchison, K. K., Spence, M. J., & Touchstone, E. W. (2008, March). Disruption of six-month-olds' infant-directed speech categorization in the presence of faces. Poster session presented at the biennial meeting of the International Conference on Infant Studies, Vancouver, BC.
Atchison, K. K., Spence, M. J., & Touchstone, E. W. (2009, April). Categorization of Synchronous Infant-Directed Speech by 4- and 6-month-old Infants. Poster session presented at the biennial meeting of the Society for Research in Child Development, Denver, CO.
Ferguson, C. (1964). Baby talk in six languages. *American Anthropologist*, New Series, Vol. 66(6), 103-114.
Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H. Papoušek, U. Jurgens, & M. Papoušek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches*. (pp. 262-282). New York: Cambridge University Press.
Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20(1), 104-113.
Katz, G. S., Cohn, J. F., & Moore, C. A. (1996). A combination of vocal fo dynamic and summary features discriminates between three pragmatic categories of infant-directed speech. *Child Development*, 67(1), 205-217.
Kuhl, P. K., & Andruski, J. E. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 227(5326), 684.
Roch, E., Mervis, C. B., Gray, D. G., Johnson, D. M., & Boyes-Braem, P. (1976). Basic Objects in Categories. *Cognitive Psychology*, 8, 382-439.
Spence, M. J., Chuang, Y., & Sokolsky, J. (2004, May). Four-Month-Old Infants Categorize Infant-Directed Speech when viewing a Female Face. Poster session presented at the biennial meeting of the International Conference for Infant Studies, Chicago, Illinois.
Spence, M. J., & Moore, D. S. (2003). Infant Directed Speech: Development from 4-6 Months. *Developmental Psychobiology*, 42(1), 97-109.
Walker-Andrews, A. S., & Bahrick, L. E. (2001). Perceiving the real world: Infants' detection of and memory for social information. *Infancy*, 2(4), 469-481.

Acknowledgments

We would like to thank the participating families for their continued support of our research. Also we would like to thank Dr. Susan Jerger for use of the drum stimulus for our control event. Additional thanks to the numerous research assistants at the Infant Learning Project.



This research was supported by a UTD Faculty Initiative Award given to Melanie J. Spence
Correspondence: Kristin Kuhlman Atchison at kkatichison@utdallas.edu or
School of Behavioral and Brain Sciences - The University of Texas at Dallas
800 W. Campbell Road MS GR 4.1; Richardson, TX 75080-3021

2009 CDS Biennial Meeting, San Antonio, TX, October 16-17, 2009

