The Role of Multisensory Information in Infant Attention to Faces of Speakers Versus the Rhythm of Speech

Lorraine E. Bahrick, Mariana Vaillant-Molina, Melissa A. Shuman, Irina Castellanos, Laura C. Batista and Mariana Werhahn

Florida International University

Abstract

According to the intersensory redundancy hypothesis (IRH), information presented redundantly and synchronously across two or more senses recruits selective attention and facilitates perceptual learning of amodal properties (e.g., rhythm, tempo, synchrony) more successfully than information that is presented to only one sense. In contrast, information available to only one sense such as the face of a silent person or a voice from a nearby room, facilitates perception of modality specific properties such as facial configuration or timbre and pitch of the voice (Bahrick & Lickliter, 2000; Bahrick, Flom, & Lickliter, 2002). Previous research has found that at the youngest ages tested, infants can detect rhythm and tempo changes in nonsocial events in bimodal, audiovisual presentations, but not in unimodal (audio or visual) presentations (Bahrick & Lickliter, 2000; Bahrick, Flom, & Lickliter, 2002). Only after a few months of additional experience with multimodal stimulation, can infants detect rhythm and tempo in both unimodal and bimodal presentations (Bahrick & Lickliter, 2004).

In this study we extended the above research to the domain of social events. We assessed whether infant sensitivity to the rhythm of speech would be facilitated in bimodal, redundant audiovisual presentations and attenuated in unimodal visual presentations, and in contrast, infant discrimination of the face of the speaker would be facilitated in unimodal visual and attenuated in bimodal, audiovisual presentations. Thirty-two 2.5-month-old infants were habituated, in an infant control procedure, to a video of one of two women speaking a nursery rhyme in one of two distinctive rhythms. They received two types of test trials, one depicting a change in the rhythm of speech and the other, a change in the face of the speaker. Visual recovery to the novel stimulus was measured. Results support the IRH and indicate that infants detect the change in the rhythm of speech only when the presentation is bimodal (audiovisual), whereas they detect the change of face only when the presentation is unimodal. These findings suggest that during social interaction, such as observing a person speaking, selective attention of the infant shifts from features of the face to temporal properties of the speech as a function of whether the visible person is speaking or silent. These results further highlight the important role of intersensory redundancy in guiding and constraining selective attention in early development.

Introduction

Infants spend much of their time in close face-to-face interactions which provide a range of multimodal and unimodal stimulation including the configuration of facial features, the pitch and timbre of the voice, and the rhythm and tempo of speech. According to the intersensory redundancy hypothesis (IRH), information presented redundantly and synchronously across two or more senses recruits selective attention and facilitates perceptual learning of amodal properties (e.g., rhythm, tempo, synchrony) more successfully than information that is presented to only one sense. In contrast, information available to only one sense such as the face of a silent person or a voice from a nearby room, facilitates perception of modality specific properties such as facial configuration or timbre and pitch of the voice (Bahrick & Lickliter, 2000; Bahrick, Flom, & Lickliter, 2002). Previous research has found that at the youngest ages tested, infants can detect rhythm and tempo changes in nonsocial events (such as a toy hammer tapping) in bimodal, audiovisual presentations, but not in unimodal (auditory or visual) presentations (Bahrick & Lickliter, 2000; Bahrick, Flom, & Lickliter, 2002). Only older infants who have had a few months of additional experience with multimodal stimulation can detect rhythm and tempo in both unimodal and bimodal presentations (Bahrick & Lickliter, 2004). In contrast, research on detection of modality specific information has shown that young infants detect a change in the face of a speaker or the orientation of a moving object in unimodal visual stimulation but not in bimodal audiovisual stimulation (Bahrick, Lickliter, Vaillant, Shuman, & Castellanos, 2004). The present study extended our research on perception of rhythm to the domain of social events (audiovisual speech), included tests for detection of both amodal information (rhythm of speech) and modality specific information (facial configuration) in the same study and for the same infants, and tested younger infants (2.5 month olds) than in our prior studies of nonsocial events. Following habituation to videos of a woman speaking, in bimodal speech, infants were expected to discriminate a change in the rhythm of speech but not the face of the speaker, whereas in unimodal visual speech infants were expected to discriminate a change in the face of the speaker but not the rhythm of speech, following predictions of the IRH.

Figure 1

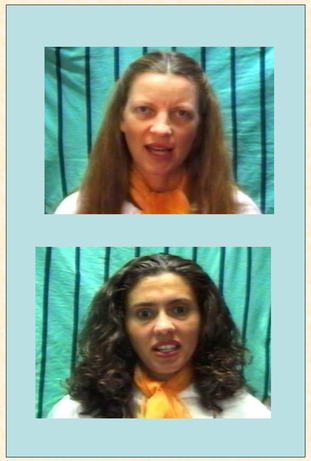
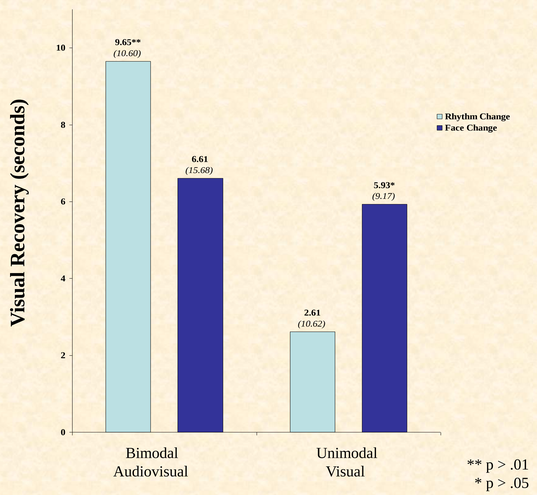


Figure 2. Mean visual recovery (and SD) to a change in rhythm of speech and face of speaker



Method

Thirty-two 2.5 month old infants were habituated, in an infant controlled procedure, to a video of one of two women (see Figure 1) speaking a nursery rhyme in one of two distinctive rhythms. Sixteen infants received bimodal audiovisual displays of the woman speaking and sixteen infants received unimodal visual displays of the woman speaking silently throughout habituation and test. Following habituation (a decrease in looking time of 50%), each infant received two types of test trials under their respective conditions. One test trial depicted a change in the rhythm of speech (with no other changes) and the other depicted a change in the woman's face (with no other changes). Visual recovery to the changes (with respect to two no-change posthabituation trials) served as the measure of discrimination.

Results

Results (see Figure 2) support our hypotheses and demonstrate that infants in the bimodal audiovisual condition showed significant visual recovery to the change in the rhythm of speech ($t(15)=3.64$, $p=.002$) but not to the change in the face of the speaker ($t(15)= 1.63$, $p>.1$). In contrast, in the unimodal visual condition infants showed the reverse pattern. They demonstrated significant visual recovery to the change in the face of the speaker ($t(15)=2.42$, $p=.031$), but not to the change in the rhythm of speech ($t(15)=.984$, $p>.1$).

Conclusions

These findings support predictions of the IRH and demonstrate that in bimodal, redundant stimulation (audiovisual speech), infants were able to discriminate a change in amodal information (rhythm of speech) but not modality specific information (facial configuration). In contrast, in unimodal visual stimulation (silent visual speech), infants were able to discriminate a change in modality specific information (facial configuration) but not amodal information (rhythm of speech). These findings suggest that during social interactions, such as observing a person speaking, selective attention of the infant shifts from features of the face to temporal properties of the speech as a function of whether the visible person is speaking or silent. These results further highlight the important role of intersensory redundancy in guiding and constraining selective attention in early development.

References

Bahrick, L.E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*, 190-201.

Bahrick, LE., Flom, R., & Lickliter, R. (2002). Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. *Developmental Psychobiology, 41*, 352-363.

Bahrick, L.E., & Lickliter, R. (2004). Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: A developmental test of the intersensory redundancy hypothesis. *Cognitive, Affective and Behavioral Neuroscience, 4*, 137-147.

Bahrick, L. E., Lickliter, R., Vaillant, M., Shuman, M., & Castellanos, I. (2004, June). The development of face perception in dynamic, multimodal events: Predictions from the intersensory redundancy hypothesis. *Fifth Annual Meeting of the International Multisensory Research Forum*, Barcelona, Spain.

Presented at the Society for Research in Child Development Biennial Meeting, April, 2005, Atlanta, GA. This research was supported by NIMH grant RO1 MH 62226 to the first author. Requests for reprints should be sent to the first author at bahrick@fiu.edu.