Journal of Experimental Child Psychology 130 (2015) 147–162

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Perception of the multisensory coherence of fluent audiovisual speech in infancy: Its emergence and the role of experience

David J. Lewkowicz [a,*], Nicholas J. Minar [b], Amy H. Tift [b], Melissa Brandon [b]

[a] Department of Communication Sciences and Disorders, Northeastern University, Boston, MA 02115, USA
[b] Department of Psychology, Florida Atlantic University, Boca Raton, FL 33431, USA

ARTICLE INFO

Article history:
Received 15 May 2014
Revised 11 October 2014
Available online 11 November 2014

Keywords:
Multisensory perception
Auditory
Visual
Speech
Infancy
Development

ABSTRACT

To investigate the developmental emergence of the perception of the multisensory coherence of native and non-native audiovisual fluent speech, we tested 4-, 8- to 10-, and 12- to 14-month-old English-learning infants. Infants first viewed two identical female faces articulating two different monologues in silence and then in the presence of an audible monologue that matched the visible articulations of one of the faces. Neither the 4-month-old nor 8- to 10-month-old infants exhibited audiovisual matching in that they did not look longer at the matching monologue. In contrast, the 12- to 14-month-old infants exhibited matching and, consistent with the emergence of perceptual expertise for the native language, perceived the multisensory coherence of native-language monologues earlier in the test trials than that of non-native language monologues. Moreover, the matching of native audible and visible speech streams observed in the 12- to 14-month-olds did not depend on audiovisual synchrony, whereas the matching of non-native audible and visible speech streams did depend on synchrony. Overall, the current findings indicate that the perception of the multisensory coherence of fluent audiovisual speech emerges late in infancy, that audiovisual synchrony cues are more important in the perception of the multisensory coherence of non-native speech than that of native audiovisual speech, and that the emergence of this skill most likely is affected by perceptual narrowing.

© 2014 Elsevier Inc. All rights reserved.

* Corresponding author.
E-mail address: d.lewkowicz@neu.edu (D.J. Lewkowicz).

http://dx.doi.org/10.1016/j.jecp.2014.10.006
0022-0965/© 2014 Elsevier Inc. All rights reserved.

Introduction

Social interactions usually involve the use of audiovisual speech (Rosenblum, 2008). Such speech consists of temporally coupled and redundant streams of audible and visible information (Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009; Munhall & Vatikiotis-Bateson, 2004; Yehia, Rubin, & Vatikiotis-Bateson, 1998). Because of its multisensory redundancy, adults usually perceive audiovisual speech as a coherent entity and not as two distinct streams of information (McGurk & MacDonald, 1976; Rosenblum, 2008; Sumby & Pollack, 1954; Summerfield, 1979; Yehia et al., 1998). This fact raises some obvious developmental questions. When in development might this ability emerge? Does it emerge in infancy? Does experience contribute to its emergence?

Several studies have investigated these questions either by asking whether infants can associate fluent audible and visible speech (Bahrick, Hernandez-Reif, & Flom, 2005; Brookes et al., 2001) or whether they can match one of two faces articulating fluent speech in two different languages with a concurrently presented audible utterance that corresponds to one of the talking faces (Dodd & Burnham, 1988; Kubicek et al., 2014; Lewkowicz & Pons, 2013). These studies have indicated that infants can associate fluent audible and visible speech and that they can match a talking face to a corresponding audible utterance, but only when the two are in infants' native language. The matching findings are especially interesting because they suggest that infants can perceive the multisensory coherence of audiovisual speech. Unfortunately, the interpretation of the latter findings is complicated by the fact that infants had access to cross-linguistic discriminative cues and that these may have facilitated audiovisual matching. If so, this raises two questions. First, can infants perceive the multisensory coherence of audiovisual speech in the absence of cross-linguistic cues? Second, if they can, at what age does this ability first emerge?

Obviously, infants should be able to perceive the multisensory coherence of fluent speech at some point—even in the absence of cross-language discriminative cues—because the perception of the multisensory coherence of their world, and especially of their native language, is fundamental to cognition (Gibson, 1969; Piaget, 1952; Rosenblum, 2008; Thelen & Smith, 1994). Most likely, however, this ability emerges relatively late in infancy for two reasons. First, speech and language perception skills emerge slowly and gradually in infancy. This is illustrated by the fact that it is not until the end of the first year of life that infants become relatively sophisticated perceivers of their native language (Saffran, Werker, & Werner, 2006; Werker, Yeung, & Yoshida, 2012). Second, multisensory processing skills also emerge slowly and gradually in infancy (Bremner, Lewkowicz, & Spence, 2012; Lewkowicz, 2014; Lewkowicz & Ghazanfar, 2009). This is illustrated by the fact that even though from birth onward infants can perceive the coherence of human auditory and visual speech (Dodd, 1979; Lewkowicz, 1996a, 2000, 2010), nonhuman communicative signals (Lewkowicz, Leo, & Simion, 2010), and nonspeech auditory and visual information (Bahrick, 1983; Brookes et al., 2001; Lewkowicz, 1986, 1992a, 1992b, 1996b), they do so only based on whether the signals in the two modalities occur together or not. It is not until the second half of the first year of life that infants begin to perceive the multisensory coherence of their audiovisual world based on more specific and more complex attributes such as gender (Patterson & Werker, 2002; Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991), affect (Walker-Andrews, 1986), and identity (Lewkowicz & Pons, 2013).

The role of audiovisual synchrony (A–V synchrony) cues in perception is especially interesting because of their fundamental importance to perception throughout the lifespan and their complex interaction with other usually concurrent multisensory relational cues. For example, some studies have found that young infants can perceive the equivalence of the facial and vocal attributes of isolated speech syllables even when the audible syllable is temporally synchronized with both visible syllables (Kuhl & Meltzoff, 1982; Patterson & Werker, 1999, 2002, 2003; Walton & Bower, 1993). This suggests that, at least in the case of single syllables, infants are able to extract phonetic multisensory invariance even in the absence of synchrony cues. Studies of older (6- and 11-month-old) infants have found similar evidence except that by then infants can even map previously heard syllables onto subsequently presented visible articulations of the same syllables (Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés, 2009).

Download English Version:

https://daneshyari.com/en/article/918008

Download Persian Version:

https://daneshyari.com/article/918008

Daneshyari.com