cortex 63 (2015) 196–205

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Sound symbolism scaffolds language development in preverbal infants

Michiko Asano [a,b,c], **Mutsumi Imai** [b,*], **Sotaro Kita** [d], **Keiichi Kitajo** [e,f], **Hiroyuki Okada** [c] **and Guillaume Thierry** [g]

[a] *Department of Psychology, College of Contemporary Psychology, Rikkyo University, Niiza, Saitama, Japan*
[b] *Faculty of Environment and Information Studies, Keio University, Fujisawa, Kanagawa, Japan*
[c] *Tamagawa University Brain Science Institute, Machida, Tokyo, Japan*
[d] *Department of Psychology, University of Warwick, Coventry, UK*
[e] *RIKEN BSI-Toyota Collaboration Center, RIKEN Brain Science Institute, Wako, Saitama, Japan*
[f] *RIKEN Brain Science Institute, Wako, Saitama, Japan*
[g] *School of Psychology, Bangor University, Gwynedd Wales, UK*

ARTICLE INFO

Article history:
Received 1 February 2014
Reviewed 10 April 2014
Revised 22 June 2014
Accepted 26 August 2014
Action editor Anne Castles
Published online 16 September 2014

Keywords:
Sound symbolism
Language development
Audio-visual correspondences
Phase synchronization analysis of EEG
Amplitude change analysis of EEG

ABSTRACT

A fundamental question in language development is how infants start to assign meaning to words. Here, using three Electroencephalogram (EEG)-based measures of brain activity, we establish that preverbal 11-month-old infants are sensitive to the non-arbitrary correspondences between language sounds and concepts, that is, to sound symbolism. In each trial, infant participants were presented with a visual stimulus (e.g., a round shape) followed by a novel spoken word that either sound-symbolically matched ("moma") or mismatched ("kipi") the shape. Amplitude increase in the gamma band showed perceptual integration of visual and auditory stimuli in the match condition within 300 msec of word onset. Furthermore, phase synchronization between electrodes at around 400 msec revealed intensified large-scale, left-hemispheric communication between brain regions in the mismatch condition as compared to the match condition, indicating heightened processing effort when integration was more demanding. Finally, event-related brain potentials showed an increased adult-like N400 response – an index of semantic integration difficulty – in the mismatch as compared to the match condition. Together, these findings suggest that 11-month-old infants spontaneously map auditory language onto visual experience by recruiting a cross-modal perceptual processing system and a nascent semantic network within the first year of life.

* *Corresponding author*. Faculty of Environment and Information Studies, Keio University, 5322 Endo, Fujisawa, Kanagawa 252-0882, Japan.
E-mail address: imai@sfc.keio.ac.jp (M. Imai).
http://dx.doi.org/10.1016/j.cortex.2014.08.025
0010-9452/

1. Introduction

How human infants map speech sounds to meaning in order to break into semantics is a key question for understanding the ontogenesis of language. It has been suggested that a biologically endowed ability to realize cross-modal mapping, particularly between auditory and visual percepts, scaffolds language learning in human infants (Imai, Kita, Nagumo, & Okada, 2008; Maurer, Pathman, & Mondloch, 2006). Consistent with this idea, 4-month-old infants appear to sense intrinsic correspondences between speech sounds and certain features of visual input (see Ozturk, Krehm, & Vouloumanos, 2013; Peña, Mehler, & Nespor, 2011), a phenomenon referred to as sound symbolism. It has also been reported that toddlers are not only sensitive to sound symbolism (Maurer et al., 2006) but also make use of sound symbolism in verb learning (Imai et al., 2008; Kantartzis, Imai, & Kita, 2011). The results from preverbal infants (Ozturk et al., 2013; Peña et al., 2011) and those from toddlers (Imai et al., 2008; Kantartzis et al., 2011; Maurer et al., 2006) support the idea that sound symbolism plays an important role in the ontogenesis of language (Imai & Kita, 2014; Imai et al., 2008; Maurer et al., 2006).

It is generally agreed that infants start to associate speech sounds and visual referents at around 12–14 months. At this age, however, the process is effortful because infants have limited information processing capacities and little experience in mapping words to the world (Fennell & Werker, 2003; Werker, Cohen, Lloyd, Casasola, & Stager, 1998). They may rely more on perceptually based cues that are available without prior word learning experiences, such as cross-modal correspondences between speech and visual input in their word learning. Indeed, previous research suggests that 14-month-old infants use sound-symbolic correspondences between speech sounds and object properties as a cue in their effort to establish word (speech sounds) – referent associations (Imai et al., under review; Miyazaki et al., 2013). Thus, there is some evidence that sound symbolism helps young infants at the initial stage of word learning.

However, how sound symbolism is processed in the infants' brain has not yet been addressed in the literature. It is not conceivable that four-month-old infants are actively engaged in semantic processing when they hear speech sounds together with a visually presented referent (Stager & Werker, 1997). Thus, infants at this age are likely to process sound symbolism perceptually, possibly on the basis of cross-modal binding mechanisms. However, at later times, the influence of sound symbolism is likely to transpire in temporal windows compatible with higher-level information processing, i.e., the semantic level.

In this study, we investigated how 11-month-old infants respond to sound symbolism. If perceptual cross-modal mapping ability scaffolds the establishment of word-referent associations, we might see the effect of sound symbolism in two time-windows: (a) in an early time window coinciding with the time period of perceptual processing, and (b) in a time window coinciding with higher-level cognition and/or semantic processing. We chose to study 11-month-olds because they are just about to say their first words but there is little or no evidence to date for the successful establishment of novel word-referent associations in experimental settings at this age. To examine this possibility, we recorded EEG from nineteen children during the presentation of novel word – visual shape pairs that were either sound-symbolically congruent or incongruent (Fig. 1) and analysed the data using three indices of brain functions: amplitude change, large-scale phase synchronization, and event-related potentials (ERPs). In each trial, infants were presented with a picture of a shape (randomly selected from 20 spiky and 20 round shapes) followed by a novel word ("kipi" or "moma").

Here, we were interested in testing whether infants would manifest increased N400 amplitude in the case of sound-symbolically mismatching word-shape pairs as compared to sound-symbolically matched ones. The N400 effect is an ERP modulation known to be sensitive to semantic integration processes in adults (Kutas & Federmeier, 2011), but also in infants (Friedrich & Friederici, 2005, 2011; Parise & Csibra, 2012). A more negative-going N400 deflection for sound symbolically mismatching sound-shape pairs would indicate that infants with very little vocabulary assume sound symbolic correspondence between word sound and shape, and consider sound-shape mismatches to be anomalies at a conceptual/semantic level.

Accumulating evidence suggests that an increase in gamma-band EEG amplitude, or gamma-band activity, is related to cross-modal perceptual integration. For example, Schneider, Debener, Oostenveld, and Engel (2008) reported that gamma-band activity increased for matched audio-visual stimuli at around 100–200 msec in the 40–50 Hz frequency range in adults (see also Senkowski, Schneider, Foxe, & Engel, 2008 for a review).

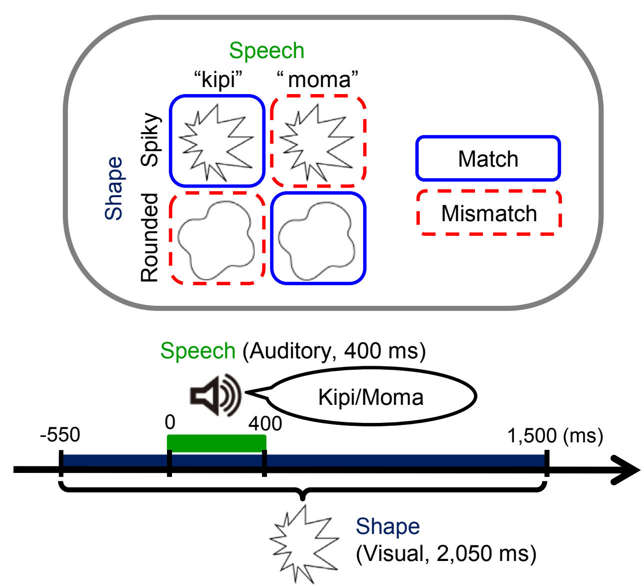

Fig. 1 – Experimental protocol. In each trial, infants were presented with a spiky or round visual shape, followed by a novel word consisting of either voiceless stops and high vowels or nasal consonants and mid/low vowels ("kipi" and "moma", respectively). Shape and sound were sound-symbolically matched (e.g., a spiky shape followed by "kipi") or mismatched (e.g., a spiky shape followed by "moma").