Experimental Neurology 250 (2013) 313–320

Contents lists available at ScienceDirect

Experimental Neurology

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

Rapid maturation of voice and linguistic processing systems in preschool children: A near-infrared spectroscopic study

Takao Yamasaki [a,b,*], Katsuya Ogata [a], Toshihiko Maekawa [a], Ikue Ijichi [a], Masatoshi Katagiri [c,d], Takako Mitsudo [a], Yoko Kamio [c], Shozo Tobimatsu [a]

[a] Department of Clinical Neurophysiology, Neurological Institute, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
[b] Department of Neurology, Minkodo Minohara Hospital, Fukuoka, Japan
[c] Department of Child and Adolescent Mental Health, National Institute of Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan
[d] Department of Neuropsychology, Graduate School of Medicine, University of Toyama, Toyama, Japan

ARTICLE INFO

Article history:
Received 23 August 2013
Revised 27 September 2013
Accepted 2 October 2013
Available online 8 October 2013

Keywords:
Voice and linguistic systems
Development
Mother's voice
Preschool children
Near-infrared spectroscopy (NIRS)

ABSTRACT

To better understand how voice and linguistic processing systems develop during the preschool years, changes in cerebral oxygenation were measured bilaterally from temporal areas using multi-channel near-infrared spectroscopy (NIRS). NIRS was recorded while children listened to their mothers' voice (MV), an unfamiliar female voice (UV) and environmental sound (ES) stimuli. Twenty typical children (aged 3–6 years) were divided into younger (Y) ($n = 10$, male $= 5$; aged 3–4.5 years) and older (O) ($n = 10$, male $= 5$; aged 4.5–6 years) groups. In the Y group, while MV stimuli significantly activated anterior temporal areas with a right predominance compared to ES stimuli, they significantly activated left mid-temporal areas compared to UV stimuli. These temporal activations were significantly higher in the Y group compared to the O group. Furthermore, only the O group exhibited significant habituation and gender differences in the left mid-temporal area during MV perception. These findings suggest that the right voice-related and the left language-related temporal areas already exist in the Y group, and that MV stimuli modulate these areas differently in the two age groups. Therefore, we conclude that a mother's voice plays an important role in the maturation of the voice and linguistic processing systems, particularly during the first half of the preschool-aged period. This role may decrease during the latter half of the preschool-aged period due to rapid development of these systems as children age.

Introduction

Voice perception and speech perception play fundamental roles in social interactions. Functional magnetic resonance imaging (fMRI) studies with adult subjects show that the superior temporal gyrus (STG) and the proximity of the superior temporal sulcus (STS) are important bilaterally in processing vocalizations, with the right hemisphere playing a dominant role (Belin and Grosbras, 2010; Belin et al., 2000). In contrast, the left-lateralized network of the STG and the left angular gyrus are responsible for language processing (Dehaene-Lambertz et al., 2002). An event-related potential (ERP) study has shown that a specific right-lateralized temporal response to voices is evident in 4- to 5-year-old children, whereas a specific left-lateralized fronto-temporal response can be observed for speech (Rogier et al., 2010). Although similar brain networks are involved in voice and language processing for both adults and children, language abilities rapidly and dramatically develop during the preschool years (3 to 6 years old) (Landers, 1990). For example, while three-year-olds have a limited vocabulary, by 5 years children can typically produce sentences that have an adult-like word order. The neural substrates for voice and linguistic processing during the first half (3–4.5 years) and the latter half (4.5–6 years) of the preschool-age period are thus likely to function differently.

Evidence shows that infants can recognize their mothers' voices shortly after birth (DeCasper and Fifer, 1980; Mehler et al., 1978). fMRI (2-month-old infants) and ERP (4-month-old infants) studies have demonstrated that infant brains respond differently to their own mothers' voices than to unfamiliar voices (Dehaene-Lambertz et al., 2010; Purhonen et al., 2004, 2005). This enhanced response to one's mother's voice from such an early age suggests that a mother's voice is especially significant for language development in typical infants (Dehaene-Lambertz et al., 2010; Liu et al., 2003). Interestingly, characteristics of maternal speech change as children age, suggesting that mothers adjust speech directed toward their children as a function of their language ability (Liu et al., 2009). Whether the special role of a mother's voice in voice processing and linguistic development is altered during the preschool-aged period is currently unknown.

Brain activity in response to sensory stimuli can be reduced by repeated exposure to the stimuli (habituation effect). This has been observed in the temporal linguistic system of adults but not infants during

* Corresponding author at: Department of Clinical Neurophysiology, Neurological Institute, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. Fax: +81 92 642 5545.
E-mail address: yamasa@neurophy.med.kyushu-u.ac.jp (T. Yamasaki).

0014-4886/$ – see front matter
http://dx.doi.org/10.1016/j.expneurol.2013.10.005

repetitive speech perception (Dehaene-Lambertz et al., 2006). A large number of studies have reported gender differences in the language abilities of children as well as adults, although these gender-related differences remain controversial (Wallentin, 2009). Therefore, investigating whether the habituation effect and gender differences exist in preschool children may advance the understanding of voice and linguistic processing system development.

Because preschool-aged children are typically very active, recording fMRI can be extremely difficult due to poor cooperation. Alternatively, near-infrared spectroscopy (NIRS) is a noninvasive neuroimaging method that measures changes in oxyhemoglobin (oxy-Hb), deoxyhemoglobin (deoxy-Hb), and total hemoglobin (total-Hb) concentrations (Hoshi et al., 2001; Villringer and Chance, 1997). The greatest advantage of NIRS over fMRI is that recordings can be made without having to fix the subject's body and head in the apparatus. NIRS is thus more suitable for brain imaging studies of preschool children. To answer our questions, changes in cerebral oxygenation were therefore measured using multichannel NIRS in typically developing preschool children. Children were segregated into two groups by age (see Participants section), and imaging was conducted under three stimulus conditions: mother's voice (MV), unfamiliar female voice (UV), and environmental sounds (ES). We hypothesized that 1) the right temporal system for voices and the left for linguistics are already present in preschool children, 2) activity patterns of these systems change during the preschool-aged period, 3) the specific changes include reduced activity in response to MV during the latter half of preschool periods due to rapid maturation, and 4) habituation and gender-related differences might appear during the preschool years.

Materials and methods

Participants

Twenty typically developing preschool children participated in this study. They were divided into younger (Y, 3–4.5 year) and older (O, 4.5–6 year) groups. Both groups included five boys and five girls (Y group: mean 48.9 ± 5.2 months; boys, 47.4 ± 6.8 months; girls, 50.4 ± 3.1 months; O group: mean 66.7 ± 4.4 months; boys, 66.0 ± 5.7 months; girls, 67.4 ± 3.0 months). All participants were right-handed and had no history of behavioral, language, or hearing dysfunction. Informed consent was obtained after the nature of the experiment had been fully explained to the mother of each child. The experimental procedures were approved by the ethics committee of the Graduate School of Medical Sciences, Kyushu University.

Stimuli

The MV, UV and ES stimuli were recorded (sampling rate, 100 kHz; resolution, 16 bit) using a WE7000 PC-based data acquisition system (Yokogawa Electric Corporation, Tokyo, Japan). For the MV, the mother of each participant read four short sentences aloud, derived from sutra, drama, nursery rhymes, and a traditional tale. Each sentence was recorded for 10 s, producing one 40-s stimulus for each child (20 total MV stimuli). The UV stimulus was recorded in the same way by one of the authors who was similar in age to the mothers in the study. This female author had no contact with any participant before or during the experiment. The ES stimulus consisting of four types of familiar sounds (rain and wind, fire-engine sirens, an ambulance, and a train) was played on a CD player. Each type of sound was recorded for 10 s to produce a 40-s ES stimulus. The mean fundamental frequency (F0) for each stimulus was determined with Praat software (http://www.praat.org Boersma, 2001), and no substantial differences were found among the stimuli (MV, 230.3 ± 39.6 Hz; UV, 234.5 ± 46.0 Hz; ES, 240.5 ± 156.7 Hz), although the ES stimuli were more variable. Mean stimulus intensities among the stimuli were adjusted to the same sound pressure level (68 dB) using customized software (Voice Evaluation System, eAdaptor Corporation, Hiroshima, Japan). Presenting auditory stimuli at this sound pressure level is comfortable for children. The three listening conditions (MV, UV, and ES) were repeated three times (40 s each) in a pseudorandom order with an intervening rest (R) condition (40 s) between each presentation using customized software (Voice Evaluation System, eAdaptor Corporation, Hiroshima, Japan; Fig. 1A) for a 12-min experiment, and counterbalanced across participants. Auditory stimuli were passively presented by a loud speaker located behind the participants. To reduce motion artifacts, participants sat on a comfortable chair and watched soundless movies (Pingu, Sony Creative Products Inc.) on a notebook PC monitor, which was placed in front of the child during NIRS recording (Fig. 1B).

NIRS recording

NIRS measurements were conducted with an ETG-4000 Optical Topography System (Hitachi Medical Corporation, Tokyo, Japan) using two 22-channel arrays of optrodes covering the left and right temporal areas (12×6 cm each; inter-optrode distance, 30 mm). On both scalp sides, the inferior midportion of the arrays was located at T3 or T4 according to the international 10–20 system for recording electroencephalograms (EEGs; Jasper, 1958; Fig. 1C). T3 and T4 positions correspond approximately to the middle temporal gyrus (MTG) and STG, including the STS (Homan et al., 1987; Okamoto et al., 2004; Steinmetz et al., 1989). Each array consisted of eight light emitters (semiconductor laser) and seven highly sensitive photodetectors (Avalanche photodiodes) that were attached to the subject's scalp via a soft silicon holder. The laser diodes emitted two principal wavelengths of light, modulated at a particular frequency for each wavelength and channel (695 ± 20 and 830 ± 20 nm, respectively; optical intensity, 2.0 mW/wavelength). The detector signals passed through a switching circuit set according to the configuration of the measuring probes and were separated by a lock-in amplifier set to the modulation frequency of the light source. These signals were measured simultaneously (sampling rate, 10 Hz), analyzed, and transformed according to their wavelengths and location using a modified Beer–Lambert equation (Cope and Delpy, 1988). This resulted in estimated changes in oxy-Hb, deoxy-Hb, and total-Hb concentration for every NIRS channel.

Data analysis

The moving-average method was employed with a 5 s window to smooth out short-term motion artifacts. An increase in neural activation typically produces an increase in oxy-Hb and a slight decrease in deoxy-Hb (Obrig and Villringer, 2003; Villringer and Chance, 1997). First we evaluated whether typical activation patterns occurred in each subject by using data averaged over multiple repetitions of the same trial type in one sequence in each participant. To obtain the averaged waveforms, the integral-mode method was employed (Nakahachi et al., 2008; Yamanaka et al., 2010). Pre- and post-activation baseline periods were defined as the 5 s before and after the activation period. Linear fitting was then applied to the data between these two baselines. Waveforms in this analysis do not reflect the actual time course of Hb concentrations, but are informative for visually evaluating whether waveforms are typical or atypical (Fig. 2). We therefore used this method to confirm typical activation patterns.

Next, waveforms representing the actual time courses of measured Hb concentration throughout the entire experiment were analyzed by the continuous mode method. We analyzed concentration changes in oxy-Hb because it is the most sensitive parameter of regional cerebral blood flow and has been found to exhibit the strongest correlation with the BOLD signal out of the three NIRS parameters (Hoshi et al., 2001; Strangman et al., 2002). Three samples of 50-s epochs (from −10 to 40 s) were averaged for each stimulus in each participant. Obvious brain responses after slow hemodynamic changes (from 15 to 35 s after stimulus onset) were analyzed.