1

Neuroscience xxx (2017) xxx–xxx

NEURAL CORRELATES OF VISUO-TACTILE CROSSMODAL PAIRED-ASSOCIATE LEARNING AND MEMORY IN HUMANS

PENG GUI, [a†‡] **YIXUAN KU,** [a,b†] **LEI LI,** [a] **XIAOJIN LI,** [c]
MARK BODNER, [d] **FRED A. LENZ,** [e]
LIPING WANG [a,b*‡] **AND YONG-DI ZHOU** [b,e,f]

[a] *Key laboratory of Brain Functional Genomics (MOE & STCSM), Institute of Cognitive Neuroscience, School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China*

[b] *NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai and Collaborative Innovation Center for Brain Science, Shanghai 200062, China*

[c] *Department of Electronic Engineering, East China Normal University, Shanghai 200062, China*

[d] *MIND Research Institute, Irvine, CA 92617, USA*

[e] *Department of Neurosurgery, School of Medicine, Johns Hopkins University, Baltimore, MD 21287, USA*

[f] *Krieger Mind/Brain Institute, Johns Hopkins University, Baltimore, MD 21218, USA*

Abstract—Studies have indicated that a cortical sensory system is capable of processing information from different sensory modalities. However, it still remains unclear when and how a cortical system integrates and retains information across sensory modalities during learning. Here we investigated the neural dynamics underlying crossmodal associations and memory by recording event-related potentials (ERPs) when human participants performed visuo-tactile (crossmodal) and visuo-visual (unimodal) paired-associate (PA) learning tasks. In a trial of the tasks, the participants were required to explore and learn the relationship (paired or non-paired) between two successive stimuli. EEG recordings revealed dynamic ERP changes during participants' learning of paired-associations. Specifically, (1) the frontal N400 component showed learning-related changes in both unimodal and crossmodal tasks but did not show any significant difference between these two tasks, while the central P400 displayed both learning changes and task differences; (2) a late posterior negative slow wave (LPN) showed the learning effect only in the crossmodal task; (3) alpha-band oscillations appeared to be involved in crossmodal working memory. Additional behavioral experiments suggested that these ERP components were not relevant to the participants' familiarity with stimuli per se. Further, by shortening the delay length (from 1300 ms to 400 ms or 200 ms) between the first and second stimulus in the crossmodal task, declines in participants' task performance were observed accordingly. Taken together, these results provide insights into the cortical plasticity (induced by PA learning) of neural networks involved in crossmodal associations in working memory. © 2017 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: crossmodal, ERP, paired-associate learning, visuo-tactile, working memory.

*Corresponding author. Address: Old Library 303, Institute of Cognitive Neuroscience, East China Normal University, 3663 N. Zhongshan Road, Shanghai 200062, China.

E-mail addresses: lipingwng@gmail.com, liping.wang@ion.ac.cn (L. Wang).

† Peng Gui and Yixuan Ku contributed equally to this work.

‡ Present address: Institute of Neuroscience, Chinese Academy of Sciences, 320 Yue Yang Road, 200031 Shanghai, China.

Abbreviations: ANOVA, analysis of variance; CI, crossmodal learning I session; CII, crossmodal learning II session; CL, crossmodal learned session; CNV, contingent negative variation; EEG, electroencephalogram; EOG, electro-oculogram; ERP(s), event-related potential(s); fMRI, functional magnetic resonance imaging; LPC, late positive component; LPN, late posterior negative slow wave; PA, paired-associate; ROI(s), region(s) of interest; RT, reaction time; S1, stimulus-1; S2, stimulus-2; SOA, stimulus onset asynchrony; UI, unimodal learning I session; UII, unimodal learning II session; UL, unimodal learned session; VT, visuo-tactile; VV, visuo-visual.

INTRODUCTION

Formation of associations between items is a crucial human ability to adapt to environmental dynamics. These associations can be quickly acquired by human beings through either explicit or implicit learning, within or across sensory modalities (Miyashita and Hayashi, 2000). In crossmodal associative learning and memory, it is essential that information about an object can be transferred from one cortical sensory system to another via cortical associations (Calvert, 2001; Fuster, 2001; Bavelier and Neville, 2002). Numerous studies in humans and non-human primates have shown that cortical neural networks responsible for crossmodal associations and memory consist of neurons from different cortical areas, including both association cortices and ''modality-specific'' sensory regions (Sakai and Miyashita, 1991; Watanabe, 1992; Gibson and Maunsell, 1997; Fuster et al., 2000; Saito et al., 2003; Tanabe and Sadato, 2009; Kassuba et al., 2013; Pillai et al., 2013; Zhang et al., 2014; Ku et al., 2015; Wang et al., 2015). Zhou and Fuster (1997, 2000) have shown that cells in the monkey primary somatosensory cortex are part of neural networks that are essential for visuo-tactile crossmodal associations and working memory. A functional magnetic imaging (fMRI) study in humans has also clarified neural substrates underlying auditory-visual crossmodal associative learning (Tanabe et al., 2005), in which, participants were required to identify predefined audio-visual or visuo-visual pairs of stimuli through the learning process based on

trial-by-trial feedback in delayed matching-to-sample tasks. Several areas, including the occipitotemporal junction and parahippocampal gyrus, showed an increase in the fMRI signal as learning proceeded only in the audio-visual crossmodal task, suggesting that widely distributed neural networks comprising both primary sensory and association cortices were involved in crossmodal associations and memory.

Earlier electroencephalographic (EEG) studies have shown ERP (event-related potential) components associated with retention and retrieval of stimulus information in unimodal (visuo-visual) paired-associate (PA) learning paradigms (Peters et al., 1977; Honda et al., 1996; Rose et al., 2001). A posterior positive component (390–1100 ms after the onset of the first stimulus) and a late frontal negative component (1100–2000 ms) have been identified, and they are assumed to be related to neural activity underlying retrieval and maintenance of PA memory respectively (Honda et al., 1996). In our previous studies, the EEG was recorded from participants performing tactile-visual or tactile-auditory crossmodal delayed matching-to-sample tasks (Ohara et al., 2006, 2008; Ku et al., 2007). The participants were required to respond to a second stimulus in the tasks immediately after its onset, indicating whether this stimulus matched the first one (sample). This pair of stimuli was preassigned and had been learned by the participants before the EEG recording. Sequential ERP components found in the study were two late positive components (LPC-1 and LPC-2, occurring around 300–700 ms after the onset of the first stimulus) related to crossmodal associations, and a late negative component (LNC, occurring around 1000–1500 ms) related to crossmodal working memory (Ohara et al., 2006, 2008; Ku et al., 2007). However, it still remains unknown how those cortical activities (ERPs) are modulated and correlated with neural processes during learning.

Studies have suggested that in memory tasks, the N400 or N400-like ERP component is related to memory encoding and retrieval (Friedman and Johnson, 2000; Kutas and Federmeier, 2011), and the late posterior negativity (LPN) around the Pz electrode is involved in memory retrieval as it has been observed in retrieving both the associative and the organizational color sources of verbal stimuli (Cycowicz et al., 2001; Cycowicz and Friedman, 2003; Johansson and Mecklinger, 2003; Herron, 2007; Nie et al., 2013). Studies have also reported the involvement of contingent negative variation (CNV) in PA learning and memory (Peters et al., 1977; Honda et al., 1996). In addition, alpha-band oscillation has been reported to play an important role in mental processes related to attention and memory (Klimesch et al., 2007; Jensen et al., 2012; Klimesch, 2012).

Using scalp EEG recording on human participants, the present study aimed to examine when and how brain activities representing crossmodal associations in working memory were progressively modulated during PA learning. We hypothesized that a series of ERP components (e.g. N400, LPN) reflecting sequential information processing in working memory would show dynamic changes along with the learning process, if they were truly involved in stimulus–stimulus paired associations. In addition, if the paired-associations to be learned were under the crossmodal condition, the ERP components would be different from those under the unimodal condition. To address the above hypotheses, we firstly examined EEG data recorded while participants were asked to perform two PA learning tasks, a visuo-tactile (VT) crossmodal task and a visuo-visual (VV) unimodal task. Dynamic changes of three ERP components (N400, LPN, CNV) and alpha-band oscillation during the delay period of the crossmodal PA learning task were focused. Since stimulus familiarity per se might also become enhanced along with PA learning, which could have effects on learning-related EEG components, the potential effects of familiarity were investigated behaviorally as well. Thirdly, additional behavioral tests were performed to investigate whether a set of delay lengths corresponding to the latencies of specific ERP components would have different effects on crossmodal and unimodal PA learning.

EXPERIMENTAL PROCEDURES

The study protocol was approved by the Committee on Human Research Protection at East China Normal University, and informed consent was obtained individually from all participants.

EEG recordings during visuo-tactile crossmodal and visuo-visual unimodal PA learning

Participants. Twenty-nine healthy volunteers were recruited for this experiment (11 males and 18 females; mean age, 22.1 years; range, 19–26 years). All of them were right-handed with normal or corrected-to-normal visual acuity, according to the Edinburgh handedness inventory (Oldfield, 1971) and E chart. Nine additional participants were recruited in the pre-experiment group to test whether stimuli and experimental procedures were appropriate for the study (without recording EEG).

Stimuli. Tests were conducted in a sound-attenuated, dimly illuminated chamber. Twelve slides of different amorphous texture patterns were used as visual stimuli, and four different frequencies of tactile vibrations were used as tactile stimuli (Fig. 1A, B). The visual stimuli were downloaded from the Internet and modified so that they were of identical size (256×256 pixels) and black-white contrast. These patterned pictures were also used in our previous study (Ku et al., 2015). A 17-inch CRT display (IBM C220P CRT; resolution ratio = 800×600 pixels; refresh rate = 60 frames per second) was used to present the visual stimuli. Participants sat on a chair facing the CRT display that was situated 1 m away. The visual stimuli were displayed in the central area of the screen at eye level and were within 5° of visual angle. Based on equal sensation contours and just noticeable difference for vibrations (Goff, 1967; Pongrac, 2008), frequencies of vibrotactile stimuli were set at 30, 80, 180, and 300 Hz. Vibrations were delivered to the tip of each participant's left index finger by a permanent magnetic