NeuroImage 52 (2010) 371–378

Contents lists available at ScienceDirect

NeuroImage

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

The “visual word form area” is involved in successful memory encoding of both words and faces

Leilei Mei [a,b], Gui Xue [c], Chuansheng Chen [b,*], Feng Xue [a], Mingxia Zhang [a], Qi Dong [a,*]

[a] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100875, China
[b] Department of Psychology and Social Behavior, University of California, Irvine, CA 92697, USA
[c] Department of Psychology, University of Southern California, Los Angeles, CA 90089, USA

ARTICLE INFO

Article history:
Received 14 December 2009
Revised 29 January 2010
Accepted 4 March 2010
Available online 29 March 2010

Keywords:
VWFA
Subsequent memory
Face
Visual word
fMRI

ABSTRACT

Previous studies have identified the critical role of the left fusiform cortex in visual word form processing, learning, and memory. However, this so-called visual word form area's (VWFA) other functions are not clear. In this study, we used fMRI and the subsequent memory paradigm to examine whether the putative VWFA was involved in the processing and successful memory encoding of faces as well as words. Twenty-two native Chinese speakers were recruited to memorize the visual forms of faces and Chinese words. Episodic memory for the studied material was tested 3 h after the scan with a recognition test. The fusiform face area (FFA) and the VWFA were functionally defined using separate localizer tasks. We found that, both within and across subjects, stronger activity in the VWFA was associated with better recognition memory of both words and faces. Furthermore, activation in the VWFA did not differ significantly during the encoding of faces and words. Our results revealed the important role of the so-called VWFA in face processing and memory and supported the view that the left mid-fusiform cortex plays a general role in the successful processing and memory of different types of visual objects (i.e., not limited to visual word forms).

Introduction

Previous studies have revealed the critical role of the left fusiform cortex in reading. First, functional imaging studies have observed strong activation in the left fusiform cortex when comparing words with nonwords, across both alphabetic and logographic writings (Cohen et al., 2000, 2002; Liu et al., 2008). Second, better reading skills are associated with greater involvement of the left fusiform gyrus (Brem et al., 2006; Schlaggar and McCandliss, 2007; Turkeltaub et al., 2003). Third, dyslexics on the other hand showed abnormal fusiform function compared to their normal counterparts (McCrory et al., 2005; Shaywitz et al., 2002; van der Mark et al., 2009). Fourth, evidence from lesion studies has revealed that damages to the left fusiform cortex (Gaillard et al., 2006) or its neural connections to other areas (Cohen et al., 2004; Epelbaum et al., 2008) resulted in impaired letter-by-letter reading.

Using the subsequent memory paradigm (i.e., comparing encoding-related brain activities of subsequently remembered and forgotten items) (Brewer et al., 1998; Wagner et al., 1998) and the training paradigm, recent research further showed the crucial role of the left fusiform gyrus in memory and learning of visual word forms. For instance, several studies using the subsequent memory paradigm have revealed that strong activation in the fusiform cortex was associated with successful encoding of both familiar words (Otten et al., 2001, 2002; Otten and Rugg, 2001; Wagner et al., 1998) and novel writings (Xue et al., submitted for publication-a). In addition, evidence from artificial language training studies has suggested that the left fusiform is optimal for learning novel visual word forms (Chen et al., 2007; Dong et al., 2008; Xue et al., 2006a). Specifically, it has been found that stronger leftward laterality of the fusiform cortex when initially processing a novel writing (pre-training) was associated with better orthographic learning after two weeks' training (Xue et al., 2006a).

Although existing studies have identified the critical role of the left mid-fusiform in learning to read, it is less clear whether this brain area is specialized for visual word form processing or it performs other cognitive functions. According to the visual word form area (VWFA) perspective (Cohen and Dehaene, 2004; Cohen et al., 2000, 2002), the left fusiform region is specialized for visual word form processing by selectively responding to familiar words. However, other researchers (e.g., Price and Devlin, 2003; Xue et al., 2006b; Xue and Poldrack, 2007) have suggested that the VWFA is not specialized for the processing of familiar visual words because there is evidence that it is also involved in lexical processing (Hillis et al., 2005; Kronbichler et al., 2004), non-word visual objects such as faces, houses, and tools (see Price and Devlin (2003) for a review), and novel writings (Xue et al., 2006b; Xue and Poldrack, 2007).

Although research on the VWFA's involvement in the processing of objects other than visual words is accumulating, it is limited in two

* Corresponding authors. Q. Dong is to be contacted at fax: +86 10 58807615. C. Chen, fax: +1 949 824 3002.
E-mail addresses: cschen@uci.edu (C. Chen), dongqi@bnu.edu.cn (Q. Dong).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.03.067

major aspects. First, these studies typically showed activation in the left mid-fusiform gyrus, but did not actually localize the activation to the VWFA. Direct comparisons between activations by familiar words and those by objects in other categories (e.g., faces) at the VWFA would provide stronger evidence. Second, only perceptual tasks were used in those studies. Thus, it is unknown whether these activations elicited by non-word objects, if they actually fall into the putative VWFA, would carry the same functional properties beyond processing into learning and memory. As mentioned above, the VWFA's activation during the processing of words (familiar or unfamiliar) usually leads to better word learning and memory. However, it is largely unknown whether activation in the same region would result in better memory of non-word objects such as faces. Although many studies have examined the neural correlates of face memory (e.g., Golarai et al., 2007; Kuskowski and Pardo, 1999; Prince et al., 2009; Xue et al., submitted for publication-b), and some have reported activation in the left mid-fusiform region (Prince et al., 2009; Xue et al., submitted for publication-b), no studies have focused on the role of VWFA in memory of faces or directly compared it with the memory of words.

Using the fMRI and the subsequent memory paradigm (Brewer et al., 1998; Wagner et al., 1998), the present study aimed to directly examine the role of the VWFA in the memory of words and faces. By using the subsequent memory paradigm, this study extended previous research by focusing on the involvement of the left fusiform region in the memory (rather than just visual processing) of words and faces. An independent localizer task was used to define the VWFA (Baker et al., 2007) and the fusiform face area (FFA) (Grill-Spector et al., 2004; Kanwisher et al., 1997; McCarthy et al., 1997). To emphasize the encoding of visual forms, an intentional encoding task was used. As shown in previous research (Bernstein et al., 2002; Otten and Rugg, 2001), perceptual and intentional encoding tasks resulted in greater engagement of the posterior regions (e.g., the fusiform cortex) in successful encoding. Subjects were explicitly instructed to memorize the visual forms. To further encourage subjects to focus on visual forms, we added homophones and the same faces from different angles to the materials to be memorized. In this study, two specific hypotheses were tested. First, we expected to replicate previous findings of the involvement of the VWFA in successful encoding of words (Otten et al., 2001, 2002; Otten and Rugg, 2001; Wagner et al., 1998). Second, we expected that the so-called VWFA would be involved in successful encoding and memory of faces. We directly compared the activation patterns and subsequent memory effects of faces and words in the VWFA.

Methods

Subjects

Twenty-two native Chinese speakers (half males; mean age = 22.8 ± 2.8 years old, with a range from 19 to 30 years) participated in this study. All subjects had normal or corrected-to-normal vision and were strongly right-handed as judged by Snyder and Harris's handedness inventory (Snyder and Harris, 1993). None of them had a previous history of neurological or psychiatric disease. Informed written consent was obtained from the subjects before the experiment. This study was approved by the IRB of the National Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University.

Materials

Four types of stimuli, including faces, Chinese words, common objects, and scrambled images of objects, were used in the localizer tasks. Each type contained 40 items. Faces and objects were taken by the same digital camera. The subsequent memory task consisted of 132 Chinese words and 132 famous faces that were neutral in emotion expressions. Famous faces were used so their familiarity to the subjects would be similar to that of familiar words. Each type of materials was further divided into two matched groups, one for the encoding task and the other as foils in the subsequent memory task. All stimuli were presented in gray-scale and 227 × 283 pixels in size.

All Chinese words were medium- to high-frequency words (higher than 25 per million according to the Chinese word frequency dictionary) (Wang and Chang, 1985), with 4–12 strokes, and 2–3 units according to the definition by Chen et al. (1996). Visual complexity (i.e., number of strokes and units) and word frequency was strictly matched across the study words, the foils, and words used in the localizer task.

The famous faces were obtained from the internet and normalized to the same resolution, brightness, and size. These stimuli were evaluated by 11 research assistants in the laboratory before experiment to ensure they were highly familiar to Chinese subjects (i.e., no items scored less than 5 on a 6-point scale with 1 representing "never seen it before" and 6 representing "very familiar"). Familiarity level and gender of the faces were matched across the study faces and the foils.

fMRI task

The fMRI task began with a localizer scan while the subject was passively viewing the four types of stimuli (faces, Chinese words, common objects, and scrambled images of objects). The 40 images of each type of materials were repeated once in the scan. The whole scan consisted of 16 consecutive 20 s epochs (4 for each type of materials), which were separated by 14 s fixation periods. Each image was presented for 750 ms, followed by a 250 ms blank interval. To ensure that subjects were awake and attentive, they were instructed to press a key whenever they noticed an image with white frame. This happened twice per epoch. The localizer scan lasted for 9 m 42 s.

After the localizer scan, participants were scanned while being asked to intentionally encode faces and words. A mixed design was used for the encoding scan, in which 6 blocks of faces interleaved with 6 blocks of words. The order of the blocks was counterbalanced across subjects. Each block included 11 stimuli and 2 successively presented fillers (homophones in the word block and different angles of the faces in the face block). During scanning, subjects were told about the fillers and were explicitly instructed to memorize the visual forms of faces or words. Subjects were further told that homophones and faces of different angles would be added in the subsequent memory test to encourage them to focus on the visual forms. In the actual test, however, no fillers were added to simplify the design. For each trial, the stimulus was presented for 2 s, followed by a blank that randomly varied from 1 to 5 s (mean = 2 s) to improve design efficiency. To avoid the primacy and recency effects, two other fillers were separately placed at the beginning and the end of the sequence. In total, the scan included 160 trials and lasted for 10 min 34 s.

Post-scan behavioral test

Three hours after scanning, a recognition test was administered to assess subjects' memory performance. Fillers in fMRI scan were excluded in this test. Consequently, a total of 132 faces and 132 words were used. For both types of the stimuli, half of them were those used in the fMRI encoding task, whereas the other half had not been seen by the subjects during the fMRI scan. All stimuli were randomly intermixed. For each stimulus, the subjects had to decide whether they had seen it during the scan on a 6-point confidence scale, ranging from 1 (definitely new) to 6 (definitely old). Each stimulus would stay on the screen until the subjects responded. The next item would appear after a 1 s blank.