

Identity versus similarity priming for letters in left mid-fusiform cortex

E. Darcy Burgund and Jennifer K. Edwards

Department of Psychology, Rice University, Houston, Texas, USA

Correspondence to Dr E. Darcy Burgund, PhD, Department of Psychology – MS 25, Rice University, PO Box 1892, Houston, Texas 77251, USA

Tel: +1 713 348 3413; fax: +1 713 348 5221; e-mail: dburgund@rice.edu

Received 2 January 2008; accepted 20 February 2008

Using functional magnetic resonance imaging, this study examined neural responses in the Visual Word Form Area to prime – test letter pairs in which visual similarity and identity were manipulated. Results revealed the greatest priming for pairs with high visual similarity, less priming for pairs with medium similarity, and the least priming for pairs with low similarity. Moreover, when visual similarity was equated, priming magnitude did not differ for

pairs with the same letter identity compared with those with different letter identities. As such, results contrast with views of the Visual Word Form Area as supporting arbitrary (nonsimilarity based) mapping requirements of reading, and suggest a less modular perspective on visual form recognition. *NeuroReport* 19:761–764 © 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Keywords: functional MRI, fusiform, letters, priming, reading, visual word form area

Introduction

The Visual Word Form Area (VWFA) hypothesis argues that a region within left mid-fusiform cortex [approximate Brodmann Area (BA) 37] is tuned specifically to the processing of visual word forms [1–3]. In line with this hypothesis, neural activity in this region is typically greater for words and pronounceable pseudowords than for nonpronounceable letter strings of the same length [3–5], and greater for strings of letters than strings of pseudoletters [6] and other nonletter stimuli (e.g. faces, houses) [7,8]. Proponents of the VWFA hypothesis contend that reading places unique demands on the visual system that cannot be satisfied by a generic recognition device [1–3]. In particular, Cohen and Dehaene [1] note that reading requires arbitrary and nonintuitive mappings of letters to sounds and meanings, such as mapping dissimilar shapes to the same sound and meaning (e.g. 'A' and 'a') and similar shapes to different sounds and meanings (e.g. 'O' and 'Q'), and argue that only a specialized processing mechanism could accomplish these exceptional goals. Indeed, these arbitrary and nonintuitive mapping demands in reading provide the computational foundation upon which the VWFA hypothesis is based.

Two important predictions regarding left-fusiform function stem directly from this computational argument for the VWFA. First, regions within the left fusiform should exhibit equivalent priming for words/letters presented in the same and different lettercase across prime and test presentations, even when letters are visually dissimilar across uppercase and lowercase versions (e.g. 'A' and 'a'). Second, regions within the left fusiform should exhibit greater priming for letters that share the same identity across prime and test presentations (e.g. A – a) than for letters that do not share the same identity across prime and test presentations

(e.g. O – Q), regardless of visual similarity. Mixed support for these predictions has been obtained. For example, equal priming has been observed for same-primed words and different-lettercase primed words (e.g. RAGE – RAGE vs. rage – RAGE) [9,10]. However, this case-invariant priming has not been observed with letter pairs (e.g. AA – AA vs. aa – AA) [11]. Moreover, although priming in the left fusiform is greater for same-primed words (e.g. RAGE – RAGE) than for anagram-primed words (e.g. GEAR – RAGE) [9], comparisons of priming in this region for same-named letters and different-named letters with equivalent visual similarity (e.g. B – b vs. K – x) have not been performed. Therefore, although computational arguments for the VWFA are theoretically compelling, evidence supporting these arguments is weak.

This study used functional magnetic resonance imaging to test the computational arguments for the VWFA directly. The VWFA was defined functionally according to the well-established criterion of greater activation for pronounceable than nonpronounceable letter strings [3–5], and activity in this region was compared between four types of prime – test conditions in which letter identity and visual similarity were manipulated. These conditions were (i) letters with different identities and low visual similarity (DIFF-LOW; e.g. F – y), (ii) letters with different identities and medium visual similarity (DIFF-MED; e.g. K – x), (iii) letters with the same identity and medium visual similarity (SAME-MED; e.g. B – b), and (iv) letters with the same identity and high visual similarity (SAME-HIGH; e.g. C – c). Letters within each prime – test condition were selected based on the similarity ratings of Boles and Clifford [12] such that items within the DIFF-LOW condition were lower in similarity than those in the DIFF-MED, SAME-MED, and

SAME-HIGH conditions (all $P < 0.001$), and items within the DIFF-MED and SAME-MED conditions did not differ from each other in similarity ($P = 0.971$) but were lower in similarity than those in the SAME-HIGH condition (both $P < 0.001$; Table 1). If a region within the left fusiform stores letter forms in a manner that is case invariant and identity specific, as argued by Cohen, Dehaene, and colleagues [1–3,9,10], equivalent priming for SAME-MED and SAME-HIGH items should be observed in this region, and priming for these items should be greater than priming for DIFF-MED and DIFF-LOW items. In contrast, if the left fusiform is sensitive to visual similarity rather than letter identity, as suggested by others [11,13], equivalent priming for SAME-MED and DIFF-MED items should be observed in this region, and this priming should be greater than that observed for DIFF-LOW items and less than that observed for SAME-HIGH items.

Method

Participants

Twelve volunteers (seven males; mean age 21.6 ± 3.0 years) participated for payment. One male volunteer was excluded owing to equipment error; thus, data from 11 participants were analyzed. Volunteers were screened using a detailed questionnaire to ensure that they had no history of neurological or psychiatric problems. In addition, all volunteers were native English speakers, right-handed, and had normal or corrected-to-normal. Informed consent was obtained from each volunteer in accordance with the guidelines and approval of the Rice University Institutional Review Board.

Materials

Prime – test pairs of letters were formed according to the criteria for the four experimental conditions (DIFF-LOW, DIFF-MED, SAME-MED, and SAME-HIGH; see Introduction and Table 1). Twenty-six pronounceable letter strings without meaning or usage in the English language were used in the letter-string portion of the experiment. Twenty-six nonpronounceable letter strings were created from

pronounceable strings by rearranging letters within each to render them unpronounceable (e.g. FREX → RFXE). All strings were four letters in length and were presented in uppercase.

Letters were presented in Geneva font, in white against a black background. Individual letters subtended approximately 1.5° of visual angle. Letter strings subtended approximately $1.5 \times 6^\circ$. Presentations and response-time measurement were controlled by the PsyScope software package [14].

Procedure

Each trial during the priming phase of the experiment consisted of a prime and test letter. The prime letter was forward and backward masked by a '#' and followed immediately by the test letter in the following manner: forward mask (500 ms), prime letter (30 ms), backward mask (15 ms), test letter (500 ms). Volunteers were instructed to decide whether the test letter had an enclosed space (e.g. 'A') or not (e.g. 'C') and to indicate their response with a button push as quickly and accurately as possible. Volunteers pushed one button with their right (or left) hand to indicate an 'enclosed' response, and another button with their left (or right) hand, to indicate a 'not-enclosed' response. The hand (left vs. right) used to indicate an 'enclosed' response was counterbalanced across volunteers. Volunteers engaged in four practice trials to ensure that they understood the instructions before the task began.

Prime – test trials were presented at intervals of 2.5, 5, and 7.5 s (average rate of one trial per 5 s). A fixation cross (+) preceded each trial and remained on the screen between trials. Volunteers engaged in four functional magnetic resonance imaging runs, each lasting approximately 4.5 min. All prime – test pairs from each of the four experimental conditions were presented within each run in a pseudo-random order.

Letter strings were presented for 1 s each at intervals of 2.5, 5, and 7.5 s (average rate of one trial per 5 s). Strings were preceded and followed by a fixation cross, and volunteers were asked to push a button with each hand as soon as each string disappeared from the screen. Volunteers completed two runs of the letter-string task, each lasting approximately 2.5 min. Within each run, half the strings were pronounceable and half were nonpronounceable.

Table 1 Prime – test letter pairs and mean similarity by condition

DIFF-LOW	DIFF-MED	SAME-MED	SAME-HIGH
A g	D b	A a	C c
B r	G e	B b	F f
C m	J t	D d	J j
E n	K x	E e	K k
F y	M w	G g	O o
H o	N z	H h	P p
f Q	c O	m M	s S
k S	q P	n N	u U
j T	p R	q Q	v V
d W	v U	r R	w W
h X	u V	t T	x X
a Y	s Z	y Y	z Z
136	265	278	416

Letter pairs are presented with the prime on the left and the test on the right. This order was varied across functional magnetic resonance imaging runs such that each letter appeared as both prime and test.

DIFF-LOW, letters with different identities and low visual similarity; DIFF-MED, letters with different identities and medium visual similarity; SAME-MED, letters with the same identity and medium visual similarity; SAME-HIGH, letters with the same identity and high visual similarity.

Image acquisition and analysis

Magnetic resonance imaging data were acquired on a Siemens 3T Allegra scanner (Siemens, Erlangen, Germany). Anatomical images were acquired first, using a transverse MP-RAGE T1-weighted sequence (Siemens) with a voxel size of $0.5 \times 0.5 \times 1$ mm (TR=1200 ms; TE=2.93 ms; flip angle= 12°). Functional images were acquired using an echo-planar sequence (TR=2500 ms; TE=40 ms; flip angle= 90° ; voxel size= 3.5×3.5 in-plane resolution). During each functional run, sets of 26 contiguous 4-mm thick axial images were acquired parallel to the anterior-posterior commissure plane.

Data from each volunteer were preprocessed to remove noise and artifacts, including correction for movement within and across runs using a rigid-body rotation and translation algorithm [15,16]. Image slices were temporally realigned (using sinc interpolation) to the midpoint of the first slice, accounting for differences in the acquisition time

for each individual slice. Data were then resampled into 2-mm isotropic voxels and warped into a standardized atlas space [17].

Preprocessed data were analyzed based on the General Linear Model (GLM) [15,18–20], using the FIDL analysis package (<http://www.nil.wustl.edu/~fidl>). Neural signals during the six conditions of interest [prime type (4); string type (2)] were modeled in the GLM at the seven time points (i.e. image acquisitions) immediately after each stimulus onset. In addition, a factor was coded to account for the within-run linear trend (linear drift and a constant term). All effects were modeled simultaneously in the GLM for each volunteer.

Regions of interest were defined based on a voxel-wise *t*-test comparing activity for pronounceable and nonpronounceable letter strings averaged across the third and fourth time points (5–7.5-s post-stimulus onset; i.e. the peak of a typical hemodynamic response function). This averaging assumes (a) that signals conform to the typical hemodynamic response function, and (b) that meaningful differences between tasks are observed in the response function's peak. The *t* statistical image produced by this analysis was smoothed by a 4-mm radius hard sphere kernel and Monte Carlo corrected using a threshold of $Z=4.0$ (24 contiguous voxels) [21,22]. A peak (local extremum) search algorithm was used to identify the coordinates [17] of activation peaks in the corrected images. Peaks separated by less than 10 mm in each image were consolidated by coordinate averaging, and spheres (20-mm diameter) were centered on each peak.

Results

Performance

Response times and error rates were analyzed via one-way repeated-measures analyses of variance comparing the four prime conditions (DIFF-LOW vs. DIFF-MED vs. SAME-MED vs. SAME-HIGH). A main effect of prime type was observed in the response times, $F(3,10)=4.99$, $P=0.006$. Tests of the simple effects revealed longer response times for DIFF-LOW (541 ms) than DIFF-MED (530 ms) items, $t(10)=3.47$, $P=0.006$, and marginally longer response times for SAME-MED (537 ms) than SAME-HIGH (518 ms) items, $t(10)=2.16$, $P=0.057$, but no difference between DIFF-MED and SAME-MED items, $t(10)=0.689$, $P=0.507$. Error rates were low (<6%) and did not exhibit a main effect, $F(3,10)=1.32$, $P=0.288$.

Neuroimaging

Three regions emerged from the voxel-wise analysis comparing activity for pronounceable and nonpronounceable letter strings (Fig. 1). In line with the VWFA hypothesis [1–4], a region within the left mid-fusiform cortex ($x=-45$, $y=-49$, $z=-14$ in Talairach and Tournoux atlas space; [17] BA 37) exhibited greater activity for pronounceable than nonpronounceable letter strings (Fig. 1a). In addition, a region in the left posterior-fusiform cortex ($x=-30$, $y=-74$, $z=-11$; BA 19) exhibited greater activity for nonpronounceable than pronounceable letter strings (Fig. 1b), and a region in the medial frontal cortex ($x=-3$, $y=-4$, $z=61$; BA 6) exhibited greater activity for pronounceable than nonpronounceable letter strings (Fig. 1c).

The hemodynamic response functions for each of the critical prime type conditions in the left mid-fusiform region

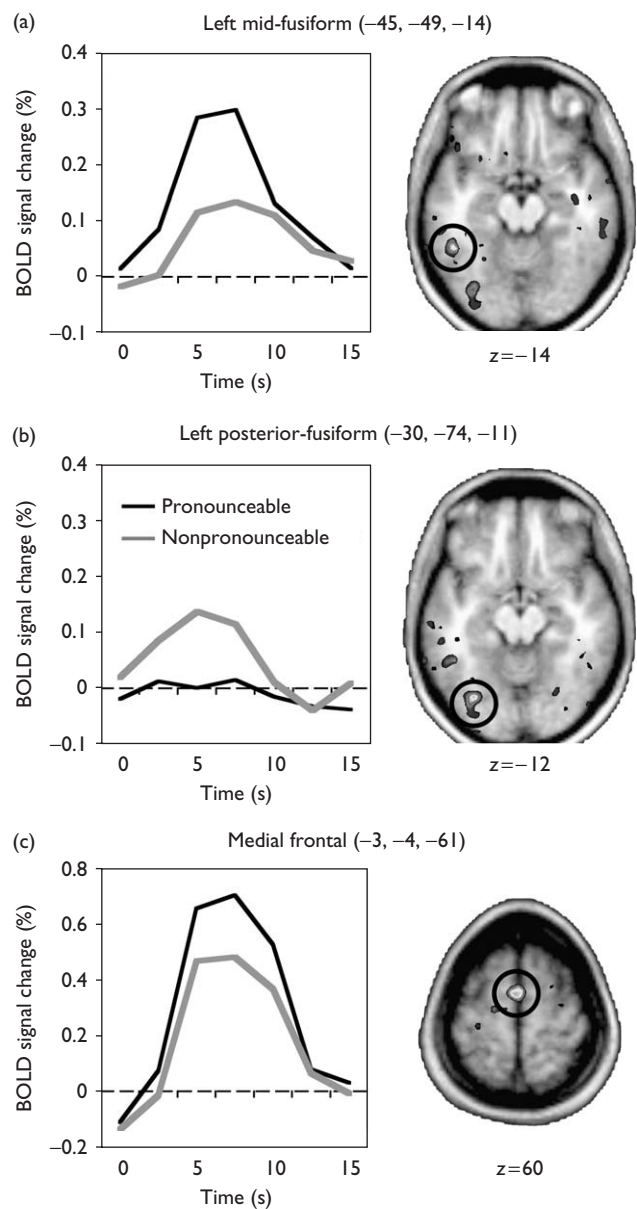


Fig. 1 Hemodynamic response in regions exhibiting an effect of letter-string type. BOLD, Blood-oxygen-level dependent.

are shown in Fig. 2. The effect of prime type was assessed in this region via a one-way repeated-measures analyses of variance comparing activity averaged across the third and fourth time points. This analysis produced a significant effect, $F(3,10)=31.63$, $P<0.001$. Tests of the simple effects revealed greater activity for DIFF-LOW (0.218) than DIFF-MED (0.109) items, $t(10)=6.46$, $P<0.001$, and greater activity for SAME-MED (0.115) than SAME-HIGH (0.049) items, $t(10)=3.80$, $P=0.004$, but no difference between DIFF-MED and SAME-MED items, $t(10)=1.08$, $P=0.307$.

Discussion

Results from this study revealed variations in priming magnitude as a function of visual similarity rather than letter identity. In particular, priming was greatest for prime – test pairs with the highest visual similarity (SAME-HIGH items),

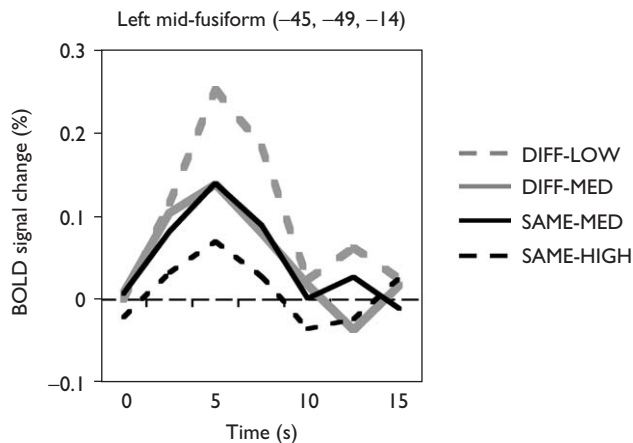


Fig. 2 Hemodynamic response in left mid-fusiform region as a function of prime type. BOLD, Blood-oxygen-level dependent; DIFF-LOW, letters with different identities and low visual similarity; DIFF-MED, letters with different identities and medium visual similarity; SAME-MED, letters with the same identity and medium visual similarity; SAME-HIGH, letters with the same identity and high visual similarity.

less for pairs with medium similarity (SAME-MED and DIFF-MED items), and least for pairs with low similarity (DIFF-LOW items). Moreover, priming magnitude did not differ for prime – test pairs with the same letter identity (SAME-MED items) compared with those with different letter identities (DIFF-MED items), when the visual similarity of the pairs was matched. Thus, results do not support predictions made by the VWFA hypothesis [1–3] and are more in line with less modular theories of visual form recognition in left mid-fusiform cortex [11,23–25].

These results converge with results from earlier studies in which letter priming in the left mid-fusiform is similarity, rather than identity, based [11,13]. However, they contrast with results from studies of word priming in which equivalent priming is observed for same and dissimilar lettercase primed words [9,10]. In contrast to single letters, whole words contain rich orthographic and semantic information that may support priming. Thus, case-invariant priming for whole words in earlier studies may have been based on this information rather than visual abstract letter representations.

Conclusion

The present work indicates that visual similarity is at least as important as letter identity in determining priming magnitude in the so-called VWFA. As such, results do not support the computational argument that this area emerges to accomplish the uniquely nonintuitive mapping requirements of reading [1]. Rather, results support a less modular perspective in which the left mid-fusiform cortex engages in similarity-driven visual object processing.

Acknowledgements

The authors thank the Brown Foundation Human Neuroimaging Lab at Baylor College of Medicine for scanner time and other resources that made the imaging possible.

References

- Cohen L, Dehaene S. Specialization within the ventral stream: the case for the visual word form area. *Neuroimage* 2004; **22**:466–476.
- Cohen L, Dehaene S, Naccache L, Lehéricy S, Dehaene-Lambertz G, Henaff MA, Michel F. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 2000; **123**:291–307.
- Cohen L, Lehéricy S, Chochon F, Lemer C, Rivaud S, Dehaene S. Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain* 2002; **125**:1054–1069.
- Dehaene S, Le Clec'H G, Poline JB, Le Bihan D, Cohen L. The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport* 2002; **13**:321–325.
- James KH, James TW, Jobard G, Wong ACN, Gauthier I. Letter processing in the visual system: different activation patterns for single letters and strings. *Cogn Affect Behav Neurosci* 2005; **5**:452–456.
- Price CJ, Wise RJS, Frackowiak RSJ. Demonstrating the implicit processing of visually presented words and pseudowords. *Cereb Cortex* 1996; **6**:62–70.
- Puce A, Allison T, Asgari M, Gore JC, McCarthy G. Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci* 1996; **16**:5205–5215.
- Hasson U, Levy I, Behrmann M, Hendler T, Malach R. Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 2002; **34**:479–490.
- Dehaene S, Jobert A, Naccache L, Ciuciu P, Poline JB, Le Bihan D, Cohen L. Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. *Psychol Sci* 2004; **15**:307–313.
- Dehaene S, Naccache L, Cohen L, Le Bihan D, Mangin J, Poline JB, Rivière D. Cerebral mechanisms of word masking and unconscious repetition priming. *Nat Neurosci* 2001; **4**:752–758.
- Burgund ED, Guo Y, Aurbach EL. Priming for letters and non-letters in fusiform cortex. Poster presented at the 37th Annual Meeting of the Society for Neuroscience, San Diego, CA: 2007.
- Boles DB, Clifford JE. An upper- and lowercase alphabetic similarity matrix, with derived generation similarity values. *Behav Res Methods Instrum Comput* 1989; **21**:579–583.
- Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW. The fusiform 'face area' is part of a network that processes faces at the individual level. *J Cogn Neurosci* 2000; **12**:495–504.
- Cohen JD, MacWhinney B, Flatt M, Provost J. PsyScope: a new graphic interactive environment for designing psychology experiments. *Behav Res Methods Instrum Comput* 1993; **25**:257–271.
- Friston KJ, Jezzard P, Turner R. Analysis of functional MRI time-series. *Hum Brain Mapp* 1994; **1**:153–171.
- Snyder AZ. Difference image vs. ratio image error function forms in PET realignment. In: Myer R, Cunningham VJ, Bailey DL, Jones T, editors. *Quantification of brain function using PET*. San Diego, CA: Academic Press; 1996. pp. 131–137.
- Talairach J, Tournoux P. *Co-planar stereotaxic atlas of the human brain*. Thieme: Stuttgart; 1988.
- Josephs O, Turner R, Friston KJ. Event-related fMRI. *Hum Brain Mapp* 1997; **5**:243–248.
- Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage* 2000; **11**:735–759.
- Zarahn E, Aguirre GK, D'Esposito M. Empirical analyses of BOLD fMRI statistics I. Spatially unsmoothed data collected under null-hypothesis conditions. *Neuroimage* 1997; **5**:179–197.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 1995; **33**:636–647.
- McAvoy MP, Ollinger JM, Buckner RL. Cluster size thresholds for assessment of significant activation in fMRI. *Neuroimage* 2001; **13**:S198.
- Joseph JE, Cerullo MA, Farley AB, Steinmetz NA, Mier CR. fMRI correlates of cortical specialization and generalization for letter processing. *Neuroimage* 2006; **32**:806–820.
- Price CJ, Devlin JT. The myth of the visual word form area. *Neuroimage* 2003; **19**:473–481.
- Price CJ, Devlin JT. The pro and cons of labeling a left occipito-temporal region the 'visual word form area'. *Neuroimage* 2004; **22**:477–479.