Title
What is the cause of left hemisphere lateralization of English visual word recognition? Pre-existing language lateralization, or task characteristics?

Permalink
https://escholarship.org/uc/item/1152n6nq

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

ISSN
1069-7977

Authors
Cottrell, Gary
Hsiao, Janet

Publication Date
2009

Peer reviewed

What is the cause of left hemisphere lateralization of English visual word recognition? Pre-existing language lateralization, or task characteristics?

Janet H. Hsiao (jhsiao@cs.ucsd.edu)
Department of Psychology, University of Hong Kong
604 Knowles Building, Pokfulam Road, Hong Kong

Garrison W. Cottrell (gary@ucsd.edu)
Department of Computer Science and Engineering, University of California San Diego
9500 Gilman Drive #0404, La Jolla, CA 92093, USA

Abstract

Visual word recognition in alphabetic languages such as English has been shown to have left hemisphere (LH) lateralization and argued to be linked to the LH superiority in language processing. Nevertheless, Chinese character recognition has been shown to be more bilateral or right hemisphere (RH) lateralized and thus is a counterexample of this claim. LH processing has been shown to have a high spatial frequency (HSF) bias, whereas RH processing has a low spatial frequency bias. Through computational modeling, here we test the hypothesis that English word recognition is lateralized to the LH and Chinese to the RH due to visual characteristics of words instead of language lateralization. We show that at least two factors may account for this dichotomy: (1) Visual similarity among words: The smaller the alphabet size is, the more similar the words in the lexicon are, and the more the model relies on HSFs to distinguish words. (2) The requirement to decompose words into letters in order to map to phonemes during learning to read English: Mapping word input to its constituent letters requires more HSF information compared with mapping to its word identity. English has a large lexicon size but only 26 letters, whereas Chinese has a much smaller lexicon with a much larger "alphabet" (stroke patterns). In addition, Chinese is a logographic system: stroke patterns do not map to phonemes and thus no decomposition is required. Hence, the lateralization of visual word recognition in different languages may depend on visual characteristics of words instead of the LH language lateralization as previously thought.

Keywords: Connectionist modeling; visual word recognition; hemispheric lateralization.

Introduction

Lateralization in face and visual word recognition

Faces are a type of visual stimuli that we have extensive exposure to from birth. The processes of face recognition have been extensively studied. For example, fMRI studies have shown that an area inside the fusiform gyrus (*fusiform face area*, FFA) responds selectively to faces, with larger activation in the right hemisphere (RH) compared with the left hemisphere (LH) (e.g. Kanwisher, McDermott, & Chun, 1997). Behaviorally, a left side bias in face perception has been consistently reported: a chimeric face made from two left half faces from the viewer's perspective (left chimeric face) is usually judged more similar to the original face than one made from two right half faces (e.g., Gilbert & Bakan, 1973). Faces elicit a larger Event Related Potential (ERP) N170 than other types of objects, especially in the RH (e.g., Rossion et al., 2003). Neuropsychological data also suggest a link between RH damage and deficits in face processing (e.g., Evans et al., 1995). In short, the RH lateralization in face processing has been consistently reported.

Words are another type of visual stimuli that we have a biased exposure to since our childhood, although slightly later than faces. In contrast to faces, the processes of visual word recognition in alphabetic languages such as English have been shown to be lateralized to the LH. Data from fMRI studies show a region inside the left fusiform area (*visual word form area*, VWFA) responding selectively to words (e.g., McCandliss, Cohen, & Dehaene, 2003). ERP studies also showed that words elicit a larger N170 in the LH than the RH (Rossion et al., 2003). A classical right visual field (RVF)/LH advantage in reading English words has been consistently reported, demonstrated first in tachistoscopic recognition (e.g., Bryden & Rainey, 1963) and subsequently in other word recognition tasks, including lexical decision (Faust, Babkoff, & Kravetz, 1995) and word naming (Brysbaert & d'Ydewalle, 1990). This phenomenon has been argued to be linked to the LH superiority in language processing, in particular phonological processing (e.g., Voyer, 1996).

In contrast, the recognition of Chinese characters, a logographic writing system, shows a left visual field (LVF)/RH advantage in tachistoscopic recognition; this effect has been argued to reflect the RH superiority in handling holistic pattern recognition (e.g., Tzeng et al., 1979; however, a recent study shows a decreased holistic processing effect in Chinese character expertise, Hsiao & Cottrell, 2009). In addition, similar to face perception, experts in Chinese character recognition have a left side bias in Chinese character perception, whereas novices do not have this bias (Hsiao & Cottrell, 2009), suggesting the RH involvement in Chinese character processing. As for phonological processing, Weekes and Zhang (1999) reported phonological priming effects on phonetic compound recognition when the characters were presented in the RVF but not LVF. In short, research in Chinese character recognition has usually exhibited a LVF advantage/bias for orthographic processing/perception and a RVF advantage for phonological processing. fMRI and ERP studies of Chinese character recognition have also exhibited

a more bilateral or RH-lateralized activation in the visual system compared with English word recognition (e.g., Tan et al., 2000; Liu & Perfetti, 2003).

Hemispheric modeling in face recognition

Why does the processing of faces, English words, and Chinese characters involve different lateralization? This is a fundamental question that goes to the heart of how and why brain areas are recruited for different tasks. Hence it is of great interest to examine how these stimuli, and the computational processes that must be carried out upon them, are similar and different. Because of the partial decussation of the optic nerves, our visual system is initially vertically split, and the two visual hemifields project contralaterally to different hemispheres. To examine at which processing stage this split information converges, we created a hemispheric model of face recognition, aiming to account for the left side bias in face perception (Hsiao, Shieh, & Cottrell, 2008). In the model we incorporated several known observations about visual anatomy and neural computation: we used Gabor responses over the input image to simulate neural responses of complex cells in the early visual system (Lades et al., 1993). We then reduced the dimension of this representation with Principal Component Analysis (PCA), a biologically plausible linear compression technique (Sanger, 1989). With this level of abstraction, convergence of the initial split may happen at three different stages: *early*, after the Gabor filters at the input layer; *intermediate*, after information extraction through PCA at the hidden layer; and *late*, at the output layer (Figure 1). The task of the model was to map each face image to its identity; in the output layer each node corresponds to a face identity (i.e. a localist representation). We also implemented a theory of hemispheric asymmetry in perception in the model, Double Filtering by Frequency (DFF, Ivry & Robertson, 1998), which argues that information coming into the brain goes through two frequency filtering stages: the first stage involves attentional selection of task-relevant frequency range, and at the second stage the LH amplifies high frequency information, whereas the RH amplifies low frequency information (cf. Monaghan & Shillcock, 2004).

Our results showed that the early convergence model failed to show the left side bias, whereas the intermediate and late convergence models showed the effect when the DFF was applied. Thus, the combination of the spatial frequency biases and the separate information extraction (PCA) is sufficient to show the left side bias, but neither alone can show the effect. This is consistent with the low spatial frequency (LSF) bias in face identification, both in humans and computational models (e.g. Schyns & Oliva, 1999; Dailey & Cottrell, 1999). The failure of the early convergence model suggests that the initially split visual input may converge at an intermediate or late stage, at which at least a certain type of information extraction has been applied separately in each hemisphere, either at or after the lateral occipital area (Hsiao et al., 2008). We also applied the model to Greeble recognition; similar to face

recognition, the results showed a left side bias in both the intermediate and late convergence models when the DFF was applied, but not in the early convergence model. This provides a testable prediction that a left side bias may also be observed in object recognition once expertise is acquired (cf. Hsiao & Cottrell, 2009).
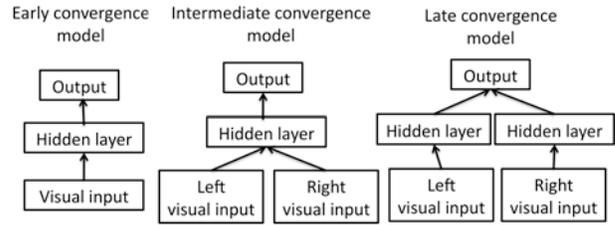

Figure 1: Hemispheric models with different timing of convergence (Hsiao et al., 2008).

Hemispheric modeling in visual word recognition

Here we extend our hemispheric model (Hsiao et al., 2008) to visual word recognition and examine under what conditions the recognition of visual stimuli relies more on high spatial frequencies/LH processing. We examine two factors that may account for the LH lateralization in English word recognition: (1) visual similarity among words in the lexicon: the more letters the words share, the more similar the words look visually; for example, to distinguish words "talk" and "walk", the visual system has to examine specifically the first letter position, and thus may have to rely more on high spatial frequencies. (2) The requirement to decompose a word into letters in order to map them into corresponding phonemes during learning to read. That is, instead of mapping the input to its word identity, the visual system has to map the word input to its constituent letter identities; this mapping may require more high spatial frequency (HSF) /fine-scale information compared with the mapping from word input to its word identity.

Maurer and McCandliss (2007) proposed the *phonological mapping hypothesis* to account for the difference in ERP N170 lateralization between faces and words: given that phonological processes are typically left-lateralized (e.g. Rumsey et al., 1997), specialized processing of visual words in visual brain areas also becomes left-lateralized. Accordingly, they pointed out that the LH lateralization in English word recognition may be related to the influence of grapheme-phoneme conversion established during learning to read, and this modulation should be less pronounced in logographic scripts such as Chinese. Note that in contrast to this hypothesis, here we do not assume the lateralization of phonological processes; we simply examine whether different mapping tasks require different spatial frequency content. In addition, previous models in visual word recognition usually use localist representations of letters (e.g. Shillcock & Monaghan, 2001) or stroke patterns (e.g. Hsiao & Shillcock, 2006a); here we bring the level of the modeling closer to the anatomy of the visual pathways in order to understand the underlying neural computation.

Modeling and Results

Here we used the same hemispheric model proposed in Hsiao et al. (2008) for face recognition, except that the input stimuli were word images instead of faces or objects (Figure 2). In the model, the input image (135 x 100 pixels) was first filtered with a grid (16 x 12) of overlapping 2D Gabor filters (Daugman, 1985) in quadrature pairs at six scales and eight orientations; the six scales corresponded to 2 to 64 cycles per word. Given the image width of 100 pixels, this frequency range can be considered as the task-relevant frequency range (the seventh scale would have 128 cycles, which exceeded the width of the image). The first stage of the DFF is implemented by simply giving this input to the model. The paired Gabor responses were combined to obtain Gabor magnitudes. Two conditions were created: in the baseline condition, Gabor responses in different scales were given equal weights; in the DFF condition, to implement the second stage of the DFF, a sigmoid-shaped weighting function was used to bias the Gabor responses on the left half word (RH) to LSFs and those on the right half word (LH) to HSFs (Figure 2). Since the early convergence model failed in accounting for the left side bias in face perception (Hsiao et al., 2008), and the late convergence model performed worse than the intermediate convergence model in the face tasks (when the DFF was applied), here we used the intermediate convergence model. Thus, in the input representation, the word image was split into left and right halves; the perceptual representation of each half was compressed into a 50-element representation (hence in total there were 100 elements); this representation can be thought of as corresponding to the lateral occipital region (i.e., structural representation of visual stimuli; analogous to Occipital Face Area in face recognition). After PCA, for each principal component, normalization was applied across all images to equalize the contribution of each component in the model. The hidden layer of the model had 20 units; it can be thought of as corresponding to the fusiform area (i.e. VWFA; analogous to FFA in face recognition).

To compare with the modeling of face recognition, here we used images of palindrome pseudo-words (such as "abccba") as the stimuli; thus the two sides of the word stimuli had the same amount of information towards the word identity (like faces). In addition, we were able to make "chimeric words" in the same way as "chimeric faces" by taking the right and left halves and making mirror images of them (see Figure 3). Thus the chimeric stimuli are different due to the original left-right orientation of the letters, and the left chimeric word matches the original on the left, but not the right due to mirror inversion, and vice versa.

For the following simulations, we ran the model 80 times and analyzed its behavior with ANOVA after 100 epochs of training (their performance on the training set all reached 100% accuracy). The training algorithm was discrete back-propagation through time (Rumelhart, Hinton & Williams, 1986), and the learning rate was 0.1. As in our previous work, after training, the size of left side bias was measured as the difference between the activation of the

output node for the original word when the left chimeric word was presented and when the right chimeric word was presented (note that output activation ranged from 0 to 1).
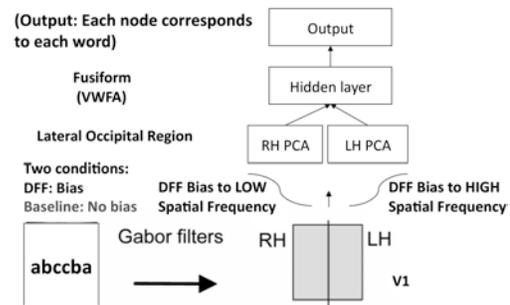

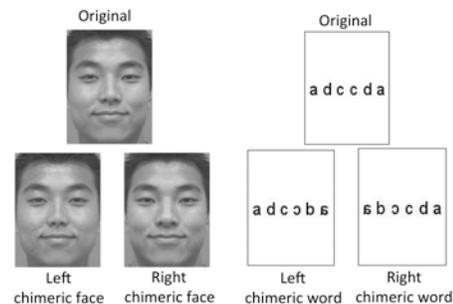Figure 2: Hemispheric model for visual word recognition.

Figure 3: Chimeric faces used in Hsiao et al. (2008), and chimeric words used in the current study.

Visual similarity among words in the lexicon

We first examined how visual similarity among words in the lexicon influences lateralization of visual word recognition. To control for the differences between English and Chinese orthographies, here we used only English letters and created artificial lexicons with different alphabet sizes. We controlled for the issue of the unbalanced nature of the location of information in a word (see Hsiao & Shillcock, 2006a) by using 6-letter palindrome words. Each lexicon had 27 words. In the first lexicon, the alphabet contained 3 letters: *a*, *b*, and *c*; thus, the 27 words in the first lexicon were all possible combinations of the three letters (*aaaaaa*, *aabbaa*, ... , *ccbbcc*, *cccccc*). In this lexicon, words were extremely similar to one another: each half of a palindrome word only differed in one letter position from that of the other palindrome words. If we increased the alphabet size to 4 (*a-d*), and randomly selected 27 words out of 64 possible words, the words in the lexicon would be less similar to one another compared with those in the first lexicon, since some words may differ in more than two letter positions. If we further increased the alphabet size and randomly selected 27 words out of all possible words, it would further decrease the similarity among words in the lexicon. Thus, the larger the alphabet size was, the less similar the words in the lexicon were, leading to a greater reliance on word shape than on letter identity. English has a large lexicon size (about 20,000 base words for a university-educated native speaker; Nation & Waring, 1997) but a small alphabet size:

only 26 letters, whereas Chinese has a much smaller lexicon (about 4,500 characters for an adult native speaker) with a much larger "alphabet" (i.e., more than 1,000 stroke patterns; Hsiao & Shillcock, 2006b). Thus, compared with Chinese characters, words in the English lexicon are visually more similar to each other. Here we started from the first lexicon, gradually increased the alphabet size to make the artificial lexicon more Chinese-like, and examined how this manipulation changed lateralization of the model.

For each simulation with a different lexicon, the materials consisted of images of 27 different words (so there were 27 output nodes). Two datasets were created: one for training and the other for testing; the order of being the training or the test set was counterbalanced across the simulation runs. In either dataset, each word had four images in different fonts (Figure 4), for a total of 108 training and 108 test images. In half of the simulation runs the mirror images of the original images were used.
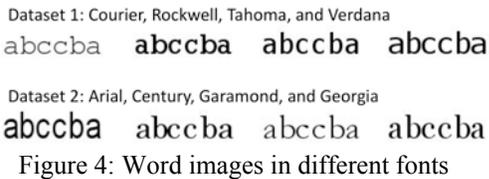
Dataset 1: Courier, Rockwell, Tahoma, and Verdana
abccba   **abccba**   abccba   abccba

Dataset 2: Arial, Century, Garamond, and Georgia
abccba   abccba   abccba   abccba

Figure 4: Word images in different fonts

Figure 5 shows the results[1]. We defined the model's RH/LSF preference as the difference in left side bias size between the DFF condition (i.e. when RH/LH were biased to LSF/HSF) and the baseline condition (i.e. when no frequency bias was applied[2]); this difference reflected the real preference over LSF/HSF biased representation due to the application of the DFF. The results showed that, when the visual similarity among the words in the lexicon was high, such as the lexicon that had only 3 letters (*a-c*), the model showed a strong LH/HSF preference ($F(1, 158) = 36.338$, $p \ll 0.001$, $\eta_p^2 = 0.187$). When we increased the alphabet size to 6 (*a-f*), the model had a weaker LH/HSF preference ($F(1, 158) = 12.653$, $p < 0.001$, $\eta_p^2 = 0.074$). When we increased the alphabet size to 7 (*a-g*) or 8 (*a-h*), this preference disappeared. When we further increased the alphabet size to 10 (*a-j*), the model started to have a weak RH/LSF preference ($F(1, 158) = 7.631$, $p < 0.01$, $\eta_p^2 = 0.046$; Figure 5). In another simulation, we further increased the alphabet size to 26 letters, the model had a strong RH/LSF preference ($F(1, 158) = 24.810$, $p \ll 0.001$, $\eta_p^2 = 0.136$). Thus, the results showed that with decreasing visual similarity among words in the lexicon, the model relied

more and more on RH/LSF information. That is, the smaller the alphabet size is, the more similar the words in the lexicon are, and the more the model relies on high spatial frequencies (i.e. LH lateralization) to distinguish words[3].
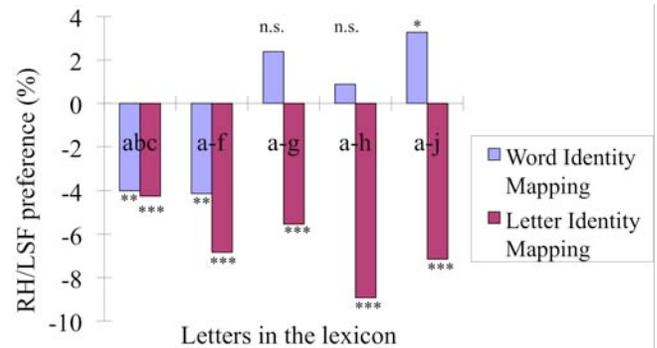

Figure 5: RH/LSF preference in the models trained with lexicons of different alphabet sizes and for different mapping tasks. In word identity mapping, it showed that the larger the alphabet size was, the more dissimilar the words in the lexicon were, and the stronger the RH/LSF preference the model had; after switching to letter identity mapping, all the models showed strong LH/HSF preference (*$p < 0.01$; **$p < 0.001$; ***$p \ll 0.001$).

Requirement to decompose a word into letters

Another major difference between English and Chinese is that there is a quasi-regular letter-sound correspondence in English, leading to phonics approaches to teaching reading, whereas there is no such correspondence in Chinese. In other words, instead of recognizing visual input of an English word as a whole, we have learned to decompose visual input of a word into its constituent letters/graphemes and then map them to corresponding phonemes when learning to read[4]. Here we explored the role of requiring the networks to map a word image to letter identities instead of to its whole-word identity. Compared with the mapping from word images to word identities, this mapping to letter identities may rely relatively more on HSF information/fine details and thus may result in LH lateralization. The same lexicons as the last section were used. To implement letter identity mapping, in the output layer, instead of having each node representing each word, we divided the output layer into three slots: the first slot corresponded to the letter identity in the first letter position (and the last since they were palindrome words), the second slot was for the letter identity in the second position, and the third slot was for the third position (Figure 6).

---

[1] Here we focus on examining the models' preference over LSF or HSF representations; we report and discuss the models' generalization performance on the test set in different conditions elsewhere. In general, the model generalized better with increasing alphabet size (Word identity mapping average accuracy: a-c: 32%; a-f: 84%; a-j: 90%; chance level was 1/27 =3.7%. Letter identity mapping average accuracy: a-c: 54%; a-f: 79%; a-j: 89%).
[2] In the baseline condition, none of the models showed any significant left or right side bias.

[3] In a separate simulation, we trained the model to recognize the 30 mirror-symmetric Chinese characters used in Hsiao & Cottrell (2009). The results showed that the model had a strong RH/LSF preference ($p \ll 0.001$). See Conclusion and Discussion.
[4] Although Chinese phonetic compounds do have a phonetic component that has partial information about the character pronunciation, this decomposition, if any, is still much coarser than that in English word recognition, since each phonetic compound has only one phonetic component.
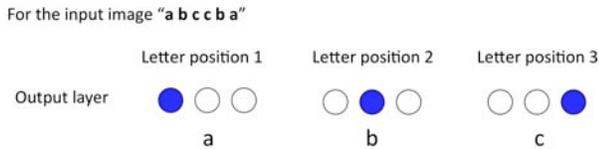
For the input image "a b c c b a"

Figure 6: Output representation for the letter identity mapping task.

The results showed that, in all simulations, the model exhibited a strong LH/HSF preference (*a-c*, $F(1, 158) = 24.486$, $p \ll 0.001$, $\eta_p^2 = 0.134$; *a-f*, $F(1, 158) = 23.516$, $p \ll 0.001$, $\eta_p^2 = 0.130$; *a-g*, $F(1, 158) = 52.353$, $p \ll 0.001$, $\eta_p^2 = 0.249$; *a-h*, $F(1, 158) = 58.945$, $p \ll 0.001$, $\eta_p^2 = 0.272$; *a-j*, $F(1, 158) = 69.274$, $p \ll 0.001$, $\eta_p^2 = 0.305$; Figure 5). The effect size measured by partial Eta squared (η_p^2) increased with the alphabet size of the lexicon. Thus, compared with the word identity task, switching the task to the letter identity mapping task significantly made the model rely more on LH/HSF information.

Conclusion and Discussion

Here we applied our hemispheric model of face/object recognition (Hsiao et al., 2008) to visual word recognition, in order to examine the factors that may account for the differences in lateralization in the processing of different languages. Visual word recognition in alphabetic languages such as English has been reported to lateralize to the LH, and argued to be due to the LH lateralization of language processes. Nevertheless, in Chinese character recognition, a RH/LVF advantage is usually reported for orthographic processing tasks (while a LH/RVF is shown for phonological tasks). This thus provides a counter example of this claim. Here we aimed to examine factors that may influence lateralization in visual word recognition other than the LH lateralization of language processes; specifically, the influence of visual characteristics of the words in a lexicon.

We first examined the influence of visual similarity among words in a lexicon. We created artificial lexicons with the same number of words of the same length, but with different alphabet sizes. We then trained the model to differentiate words in the lexicons (word identity mapping). We show that the smaller the alphabet size is, the more similar the words in the lexicon are, and the more the model relies on HSF/LH representations to distinguish words. We then examined the influence of the requirement to decompose words into letters/graphemes in order to perform grapheme-phoneme conversion during learning to read. We used the same artificial lexicons but changed the task from word identity mapping to letter identity mapping. We show that the letter identity mapping task requires more HSF information compared with the word identity mapping task. In contrast to the phonological mapping hypothesis (Maurer & McCandliss, 2007), which argues that visual word processing becomes LH lateralized due to the left-lateralized phonological processes, our model did not use/assume any phonological processes/representations; the task of the model was purely visual – to recognize constituent letters of the word input.

The two visual factors we identified here are able to account for the difference in lateralization between visual word recognition in English and Chinese. English has a large lexicon size (about 20,000 base words for a native speaker) but a small alphabet size: only 26 letters, whereas Chinese has a much smaller lexicon (about 4,500 characters for a native speaker) with a much larger "alphabet" (more than 1,000 stroke patterns). Thus, compared with Chinese characters, words in the English lexicon are visually more similar to each other; our results suggest that distinguishing English words may require more HSF information compared with distinguishing Chinese characters. In addition, Chinese is a logographic system, in which stroke patterns do not map to phonemes. Unlike English readers, Chinese readers do not rely on grapheme-phoneme conversion during learning to read, and thus there is no requirement to decompose a character into its constituent stroke patterns in order to read. Our results thus suggest that English word recognition requires more HSF information compared with Chinese character recognition due to its task characteristics during learning to read. We are currently working on the modeling with English and Chinese lexicons in a realistic scale and examining their lateralization.

Note that our results do not rule out the possibility of the influence of LH lateralized phonological processes on the lateralization of visual word recognition. In an ERP study, Hsiao, Shillcock, and Lee (2007) examined N170 amplitude elicited by characters with their phonetic component on the left or right. They showed that characters with their phonetic component on the left (which is initially projected to the RH when the character was centrally presented) elicited N170 with similar amplitude in the two hemispheres, whereas characters with their phonetic component on the right (which is initially projected to the LH) elicited larger N170 amplitude in the LH than that in the RH. This phenomenon suggests that the LH lateralized phonological processing may also influence the lateralization of visual word recognition.

In summary, here we show that the lateralization of visual word recognition in different languages may depend on visual characteristics of words instead of the LH language lateralization as previously thought; specifically, it can be influenced by (1) visual similarity among words in the lexicon, and (2) the requirement to decompose a word input into its constituent letters/graphemes for grapheme-phoneme conversion during learning to read.

References

Bryden, MP. & Rainey, CA. (1963). Left-right differences in tachistoscopic recognition. *Journal of Experimental Psychology*, *66*, 568–571.

Brysbaert, M. & d'Ydewalle, G. (1990). Tachistoscopic presentation of verbal stimuli for assessing cerebral dominance: Reliability data and some practical recommendations. *Neuropsychologia*, *28*, 443-455.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, *2*, 1160–1169.

Dailey, MN. & Cottrell, GW. (1999) Organization of face and object recognition in modular neural network models. *Neural Networks*, *12(7-8)*, 1053-1074.

Evans, JJ., Heggs, AJ., Antoun, N., & Hodges, JR. (1995). Progressive prosopagnosia associated with selective right temporal lobe atrophy: a new syndrome? *Brain*, *118*, 1-13.

Faust, M., Babkoff, H., & Kravetz, S. (1995). Linguistic processes in the two cerebral hemispheres: Implications for modularity vs interactionism. *Journal of Clinical and Experimental Neuropsychology*, *17*, 171-192.

Gilbert, C. & Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia*, *11*, 355-362.

Hsiao, JH. & Cottrell, GW. (2009). Not all expertise is holistic, but it may be leftist: The case of Chinese character recognition. *Psychological Science*, *20(4)*, 455-463.

Hsiao, JH., Shieh, D., & Cottrell, GW. (2008). Convergence of the visual field split: hemispheric modeling of face and object recognition. *Journal of Cognitive Neuroscience*, *20(12)*, 2298-2307.

Hsiao, JH. & Shillcock, R. (2006a). Hemispheric differences emerge from perceptual learning: Evidence from modeling Chinese character pronunciation. Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society (pp. 345-350).

Hsiao, JH. & Shillcock, R. (2006b). Analysis of a Chinese phonetic compound database: Implications for orthographic processing. *Journal of Psycholinguistic Research*, *35*, 405-426.

Hsiao, JH., Shillcock, R., & Lee, C. (2007). Neural correlates of foveal splitting in reading: evidence from an ERP study of Chinese character recognition. *Neuropsychologia*, *45*, 1280-1292.

Ivry, R. & Robertson, LC. (1998). *The Two Sides of Perception*. Cambridge: MIT Press.

Kanwisher, N., McDermott, J. & Chun, MM. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302-4311.

Lades, M., Vorbruggen, JC., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, RP., & Konen, W. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans. on Computers*, *42*, 300–311.

Liu, Y., & Perfetti, CA. (2003). The time course of brain activity in reading English and Chinese: An ERP study of Chinese bilinguals. *Human Brain Mapping*, *18*, 167-175.

Maurer U., & McCandliss, BD. (2007). The development of visual expertise for words: the contribution of electrophysiology. In E. L. Grigorenko & A. Naples (Eds). *Single-Word Reading: Cognitive, behavioral and biological perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.

McCandliss, BD., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*, 293-299.

Monaghan, P. & Shillcock, RC. (2004). Hemispheric asymmetries in cognitive modeling: connectionist modeling of unilateral visual neglect. *Psychological Review*, *111*, 283-308.

Nation, P & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In Schmitt N, & McCarthy, M. (Eds). *Vocabulary: Description, Acquisition, Pedagogy*. New York: Cambridge University Press, pp. 6-19.

Rossion, B., Joyce, CA., Cottrell, GW., & Tarr, MJ. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage*, *20*, 1609-1624.

Rumelhart, DE., Hinton, GE., & Williams, RJ. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536.

Rumsey, JM., Horwitz, B., Donohue, BC., Nace, K., Maisog, JM., & Andreason, P. (1997). Phonological and orthographic components of word recognition. A PET-rCBF study. *Brain, 120(5)*, 739-759.

Sanger, T. (1989). An optimality principle for unsupervised learning. In Touretzky, D. (ed), *Advances in Neural Information Processing Systems,* vol. 1, pp. 11-19, San Mateo: Morgan Kaufmann.

Schyns, PG. & Oliva A. (1999). Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243-265.

Shillcock, RC. & Monaghan, P. (2001). The computational exploration of visual word recognition in a split model. *Neural Computation*, *13*, 1171-1198.

Tan, LH., Spinks, JA., Gao, JH., Liu, HL., Perfetti, CA., Xiong, J., et al. (2000). Brain activation in the processing of Chinese characters and words: A functional MRI study. *Human Brain Mapping*, *10*, 16-27.

Tzeng, OJL., Hung, DL., Cotton, B., & Wang, SY. (1979). Visual lateralization effect in reading Chinese characters. *Nature* (London), *282*, 499-501.

Voyer, D. (1996). On the magnitude of laterality effects and sex differences in functional lateralities. *Laterality*, *1*, 51-83.

Weekes, BS., & Zhang, BY. (1999). Chinese character recognition in the left and right visual fields. *Brain & Cognition*, *40*, 269-272.