Computational exploration of the relationship between holistic processing and right hemisphere lateralization in featural and configural recognition tasks

Janet Hui-wen Hsiao (jhsiao@hku.hk)
Kloser Chee Fung Cheung (kloser@hku.hk)
Department of Psychology, University of Hong Kong
604 Knowles Building, Pokfulam Road, Hong Kong SAR

Abstract

Holistic processing has long been considered as a property of right hemisphere (RH) processing. Nevertheless, recent studies showed reduced holistic processing and increased RH lateralization in Chinese character recognition expertise, suggesting that these two effects may separate. Through computational modeling, in which we implement a theory of hemispheric asymmetry in perception that posits a low frequency bias in the RH and a high frequency bias in the left hemisphere, we show that when the recognition task relies purely on featural information, holistic processing increases whereas RH lateralization decreases with increasing stimulus similarity; there is a negative correlation between them. In contrast, when the task relies purely on configural information, although RH lateralization negatively correlates with stimulus similarity, holistic processing does not correlate with stimulus similarity; there is a positive correlation between them. This suggests that holistic processing and RH lateralization do not always go together, depending on the task requirements.

Keywords: holistic processing, hemispheric asymmetry, computational modeling.

Introduction

Holistic processing and right hemisphere lateralization

In face recognition, a holistic processing effect has been consistently reported; it refers to the phenomenon that we view faces as a whole instead of various parts, and has been argued to be related to our expertise in face processing (e.g., Bukach, Gauthier, & Tarr, 2006; although some argue that it is specific to faces; e.g., McKone, Kanwisher, & Duchaine, 2007). Subsequent studies suggest a correlation between an increase in holistic processing and expertise in subordinate-level individualization as opposed to basic-level categorization (e.g., Gauthier et al., 1998; Wong, Palmeri, & Gauthier, 2009).

In addition to holistic processing, another well-known effect in face recognition is the right hemisphere (RH) lateralization effect. For example, behaviorally a left side bias in face perception has been observed: a chimeric face made from two left half faces from the viewer's perspective is usually judged more similar to the original face than one made from two right half faces (Gilbert & Bakan, 1973); this effect has been argued to be an indication of the RH involvement in face processing (e.g., Burt & Perrett, 1997). fMRI studies show that an area inside the fusiform gyrus (*fusiform face area*) responds selectively to faces (although some argue that it is an area for expertise in subordinate-level visual processing instead; Tarr & Gauthier, 2000) with larger activation in the RH than the left hemisphere (LH) (e.g. Kanwisher, McDermott, & Chun, 1997). ERP data show that faces elicit larger N170 than other types of objects, especially in the RH (Rossion et al., 2003). Neuropsychological data also suggest a link between RH damage and deficits in face recognition (e.g. Meadows, 1974). In short, the RH lateralization in face processing has been consistently reported.

The holistic face processing effect has been shown to be linked to brain activation in face selective areas especially in the RH (e.g., Schiltz et al., 2010; Harris & Aguirre, 2008). It has also been shown that the increase in holistic processing after artificial object recognition training is correlated with right fusiform area activity (Gauthier & Tarr, 2002). These results are consistent with the hemispheric asymmetry literature that posits a holistic/analytic dichotomy between RH and LH processing (e.g., Bradshaw & Nettleton, 1981), and suggest a close relationship between holistic processing and RH lateralization. Nevertheless, Hsiao and Cottrell (2009) recently showed that Chinese character recognition experts have reduced holistic processing and increased RH lateralization in processing Chinese characters compared with novices. This effect suggests that holistic processing and RH lateralization may be separate processes that do not always go together.

Faces and Chinese characters differ in both featural and configural dimensions[1]. In the featural dimension, faces consist of common components (i.e. the eyes, nose, and mouth) and the components of different faces usually look similar to each other; in contrast, Chinese character recognition involves discriminating different combinations of more than a thousand stroke patterns (Hsiao & Shillcock, 2006), which usually look dissimilar to each other. In the configural dimension, second-order spatial relations (i.e. distances) between face components have been shown to be more important in face recognition than in the recognition of other visual object classes (e.g., Farah et al., 1998), whereas this configural information is not important in Chinese character recognition, since changes in distance among character components do not change the character identity (e.g., Ge et al., 2006). In order to understand how difference in task requirements in either the featural or the configural dimension modulates holistic processing and RH lateralization in recognition tasks, here we adopt a computational modeling approach, since modeling allows

[1] Note that in the literature of face recognition, the definition of of configural processing often varied among studies. Here we refer to the configural dimension as second-order spatial relations (e.g., Mondloch, Grand, & Maurer, 2002).

good control over variables that may be hard to tease apart in human subject studies. We introduce our model below.

Hemispheric processing model

Anatomically our visual field is split along the vertical midline, with the two hemifields initially contralaterally projected to the two hemispheres. In order to examine at which processing stage this split information converges, Hsiao, Shieh, and Cottrell (2008) conducted a modeling study aiming to account for the left side bias effect in face perception. They proposed 3 models with different timings of convergence: early, intermediate, and late (Figure 1), and showed that both the intermediate and late convergence models were able to account for the effect, whereas the early convergence model failed to show the effect.

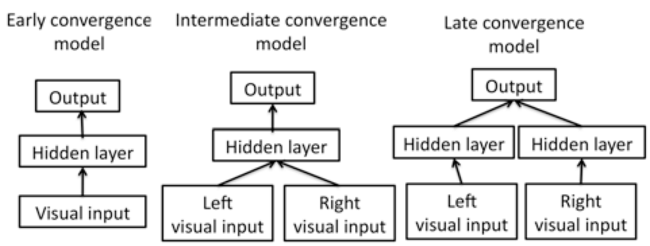

Figure 1: Hemispheric models with different timing of convergence (Hsiao et al., 2008).

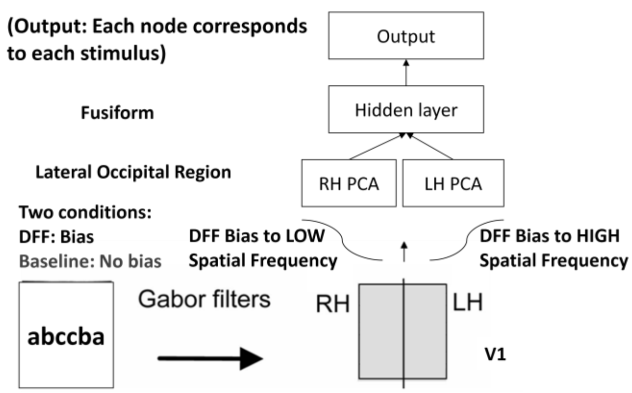

Figure 2: Hsiao et al.'s model (2008).

Hsiao et al.'s (2008) model incorporated several known observations about visual anatomy and neural computation. They used Gabor responses over the input images to simulate neural responses of cells in the early visual area (Lades et al., 1993), and Principal Component Analysis (PCA), a biologically plausible linear compression technique (Sanger, 1989), to simulate possible information extraction processes beyond the early visual area. They then used this PCA representation as the input to a two-layer neural network (Figure 2). In addition, they implemented a theory of hemispheric asymmetry in perception, Double Filtering by Frequency theory (DFF, Ivry & Robertson, 1998) in the model. The theory posits that visual information coming into the brain goes through two frequency-filtering stages: The first stage involves attentional selection of a task-relevant frequency range. At the second stage, the LH amplifies high spatial frequency (HSF) information, while the RH amplifies low spatial frequency (LSF) information. This differential frequency bias in the two hemispheres was implemented in the model by using two sigmoid functions assigning different weights to the Gabor responses in the two hemispheres (Figure 2).

Modeling holistic processing effects

In human studies, holistic processing is usually assessed through the composite paradigm (e.g., Gauthier & Bukach, 2007). In this paradigm, two stimuli are presented briefly, either sequentially or simultaneously. Participants attend to either the top or bottom halves of the stimuli and judge whether they are the same or different. In congruent trials, the attended and irrelevant halves lead to the same response, whereas in incongruent trials, they lead to different responses. Holistic processing is indicated by the interference from the irrelevant halves in matching the attended halves; it can be assessed by the performance difference between the congruent and the incongruent trials (Figure 3).

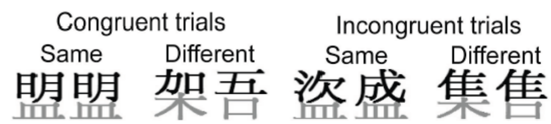

Figure 3: The complete composite design with Chinese characters as the stimuli. The example shows the condition in which the bottom halves (components in grey) are the attended halves (taken from Hsiao & Cottrell, 2009).

The holistic face processing effect has been accounted for by computational models. For example, Cottrell, Branson, and Calder (2002) trained a computational model to perform a face identification task and an expression judgment task, and showed that the model was able to account for holistic processing effects in both tasks. Richler, Mach, Gauthier, and Palmeri (2007) also used a variant of Cottrell et al.'s (2002) model to account for the holistic processing effect in face recognition. Similar to Hsiao et al.'s (2008) early convergence model (Figure 1), Richler et al.'s model (2007) applied Gabor filters to the input image, followed by PCA, and then a two-layer neural network performing the classification task, without a split architecture or frequency biases. To assess holistic processing effects, after training the hidden layer representation of each input face image was used as its internal representation in the visual working memory. Selective attention to the cued part in the composite paradigm was simulated by attenuating the Gabor response representation of the unattended half by a factor of 0.125. In each trial, the correlation between the representations of each pair of faces was used as the similarity measure; the difference in this measure between same and different trials was used to calculate d′, and the difference in d′ between the congruent and incongruent conditions was used as the measure of holistic processing. Here we apply the method used by Richler et al. (2007; cf. Cottrell et al., 2002) to assess holistic processing in our model.

Configural vs. featureal recognition tasks

To examine people's sensitivity to configural and featural changes in face recognition, Mondloch et al. (2007) created some carefully controlled datasets: in the configural set, faces had the same features (i.e., the eyes, mouth, and nose) and differed only in the distance between these features; in contrast, in the featural dataset, faces had the same distance/configuration among features but the features were different. Here we create our datasets in a similar fashion, in order to examine the relationship between holistic processing and RH lateralization when the recognition tasks depend on either configural or featural information. We use artificial stimuli that consist of three features forming a triangular configuration, and the features are taken from letters in the English alphabet. In a configural recognition task, all stimuli have the same three features ('a'), but their configurations differ (Figure 4(a)). In contrast, in a featural recognition task, all stimuli have the same configuration but the features differ (Figure 4(b)). The model is trained to recognize the stimuli in the dataset. In the configural tasks, we examine the effect of stimulus similarity in the dataset by manipulating the number of possible locations in which a feature can appear. Similarly, in the featural tasks, we examine the effect of stimulus similarity in the dataset by manipulating the number of possible letters appearing in each feature position. We aim to examine how different recognition task requirements (configural vs. featural) modulate holistic processing and hemispheric lateralization effects and the relationship between the two effects.

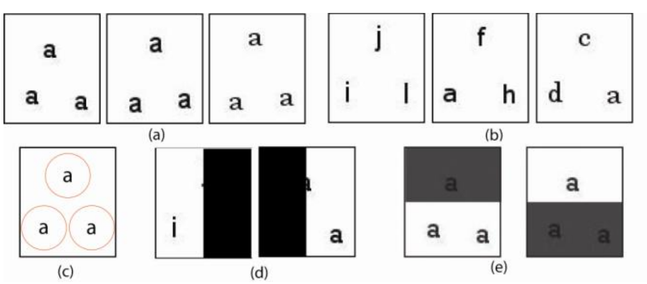

Figure 4: (a) Images in the configural set; (b) Images in the featural set; (c) Stimulus design; the three circles indicate the area of possible positions for each letter. (d) Right and left damaged images; (e) Top and bottom attenuated images.

Modeling Methods and Results

All images we used were 80x70 pixels having three English letters as features forming a triangular configuration, with one letter on the top and the other two on the bottom (Figure 4(c)). In the configural dataset, for each stimulus we fixed the identity of the letters ('a'), and the position of each letter was assigned randomly within a circular area of radius 8 pixels (Figure 4(c)). In contrast, in the featural dataset, for each stimulus we fixed the letter positions and randomly chose one letter for each position from a fixed set of letters (i.e. the English alphabet).

To create datasets with different stimulus similarities, in the configural datasets, we varied the numbers of possible locations each letter 'a' could appear within each circular area. In total we created 9 configural datasets, with the number of possible locations ranging from 4 to 12. In each dataset, 26 stimuli were randomly selected from all possible location combinations. Similarly, we varied the number of possible letters that could appear in each letter position in the featural datasets, ranged from 4 to 12, and in total 9 datasets created. In each dataset, 26 stimuli were randomly selected from all possible feature combinations. In these datasets, while keeping the total number of stimuli fixed, increasing the number of possible locations/features made the stimuli less and less similar to each other (see, e.g., Cheung & Hsiao, 2010).

In the simulations, each stimulus had 8 images, each of which had a different font. We used 4 fonts for training and the other 4 for testing (counterbalanced across simulation runs), resulting in a total number of 104 images in each of the training and testing sets. Thus, we were able to test the model's generalization ability across different fonts.

In the modeling, an input image was first filtered with a 14x12 rigid grid of overlapping 2D Gabor Filters (Daugman, 1985). At each grid point, we applied Gabor filters of 8 orientations and 5 scales (the task-relevant frequency range, depending on the image size. The maximum frequency should not exceed 2 pixels per cycle; the 6th scale, $2^6 = 64$ cycles per image exceeds the maximum frequency of the images, 70/2 = 35 cycles per image). Thus, each image was transformed into a vector of size 6,720 (14x12 sample points x 8 orientation x 5 scales). After obtaining the Gabor response representations, two conditions were created: (1) the baseline condition, in which equal weights were given to different scales of the Gabor responses; (2) the biased condition, in which we implemented the second stage of the DFF theory by using a sigmoidal weighting function to bias the responses on the left half image (RH) to LSFs, and those on the right half image (LH) to HSFs (Figure 2). The left and right perceptual representations were then compressed by PCA separately into a 50-element representation each. This representation was then used as the input to a two-layer neural network (See Hsiao et al., 2008 for more details).

We trained the model to recognize the stimuli until the performance on the training set reached an expert level (100% accuracy). In the output layer of the neural network, each output node corresponded to a stimulus identity (thus there were 26 output nodes). We used gradient descent with an adaptive learning rate as our training algorithm.

To examine hemispheric lateralization effects, after training we tested the model with images that had a damaged RH or LH PCA representation (Figure 2) by setting the PCA representation to zeros (Figure 4(d)). Thus, when mapping these damaged images to their identities, only the information from one of the visual hemifields was used. The left side (RH) bias was assessed by the accuracy difference between recognizing a right-side-damaged stimulus (carrying RH/LSF information) as the original stimulus and recognizing a left-side-damaged stimulus (carrying LH/HSF information) as the original one. We defined RH lateralization (RH/LSF preference, Hsiao et al., 2008) as the left side bias measured in the biased condition minus that measured in the baseline condition.

To examine holistic processing effects, after training we attenuated the Gabor responses of either the top or bottom half of the images in the test set by multiplying a factor of 0.125 to simulate directing the model's attention to the bottom or top half of the images respectively (Richler et al., 2007; Figure 4(e)). The complete composite design was used; it has been shown to be more robust than the partial composite paradigm (Gauthier & Bukach, 2007; Richler, Cheung, & Gauthier, in press). We created 4 types of stimulus pairs corresponding to the 4 conditions shown in Figure 3. Twenty pairs of images in each condition were randomly selected to form the materials (80 pairs in total). We calculated the correlation of the hidden layer representations in each pair as the similarity measure between them. A threshold was set to be the midpoint between the mean correlation of the "same" stimulus pairs and that of the "different" stimulus pairs. We assumed that the model responded "same" when the correlation of a pair was higher than the threshold, and responded "different" when the correlation was lower than the threshold. The holistic processing effect was indicated by the discrimination perfomance difference between the congruent and incongruent trials measured by d'.

Configural recognition tasks

The results showed that in all configural tasks, there was a significant RH lateralization effect (Figure 5(a)). Nevertheless, RH lateralization did not change significantly with the number of possible locations each letter could appear (r = 0.007, n.s.). Figure 5(b) showed the holistic processing effect (i.e. the difference between the congruent and incongruent trials) in the biased condition: holistic processing decreased as the number of possible locations increased (r = -0.209, $p < 0.001$). To further explore the relationship among stimulus similarity, RH lateralization, and holistic processing, we examined the correlations among them. We considered the Gabor responses of each stimulus as a point in a high-dimensional space; the dissimilarity among stimuli in a dataset was calculated as the average distance among these points in the space using the Unweighted Pair Group Method with Arithmetic Means (UPGMA; see Legendre & Legendre, 1998). The results showed a positive correlation between RH lateralization and stimulus dissimilarity: the less similar the stimuli were, the more RH lateralization the model exhibited (r = 0.437, $p <$ 0.001; Figure 6(a)); in contrast, there was no correlation between holistic processing and stimulus dissimilarity (r = -0.013, n.s.; Figure 6(b)). There was a weak positive correlation between holistic processing and RH lateralization (r = 0.048, $p < 0.05$; Figure 7(a)). This effect suggested that when the recognition task mainly relies on configural information, the more RH lateralization the model had, the stronger the holistic processing effect the model exhibited.

Featural recognition tasks

The results showed that in all featural tasks, there was a significant RH lateralization (Figure 8(a)); this RH lateralization increased as the number of possible letters in each letter position increased (r = 0.597, $p < 0.001$). Figure 8(b) showed the holistic processing effect in the biased condition: similar to the configural tasks, holistic processing decreased as the number of possible letters increased. Regarding the relationship among stimulus similarity, RH lateralization, and holistic processing, similar to the configural tasks, there was a positive correlation between RH lateralization and stimulus dissimilarity (r = 0.600, $p <$ 0.001; Figure 9(a)). In contrast to the configural tasks, there was a negative correlation between holistic processing and stimulus dissimilarity: the more similar the stimuli were, the stronger the holistic processing was (r = -0.256, $p < 0.001$; Figure 9(b)); in addition, there was a negative correlation between holistic processing and RH lateralization: the weaker the holistic processing was, the stronger the RH lateralization was (r = -0.211, $p < 0.001$; Figure 7(b)).

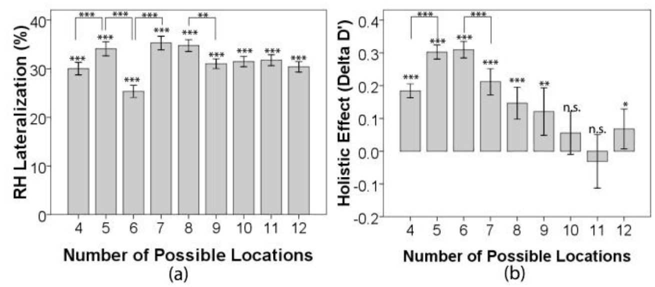

Figure 5: Configural tasks: (a) RH lateralization, (b) Holistic processing in the biased condition (comparisons with 0 and pair comparisons, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

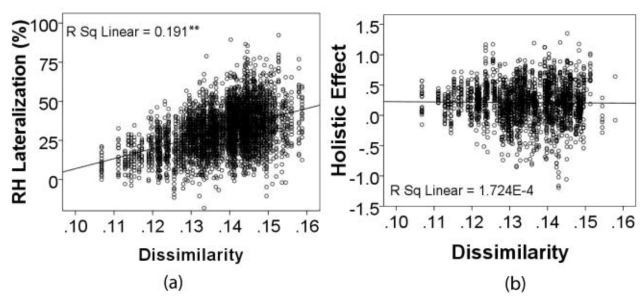

Figure 6: Configural tasks: Relationship between (a) stimulus dissimilarity and RH lateralization, and (b) stimulus dissimilarity and holistic processing.

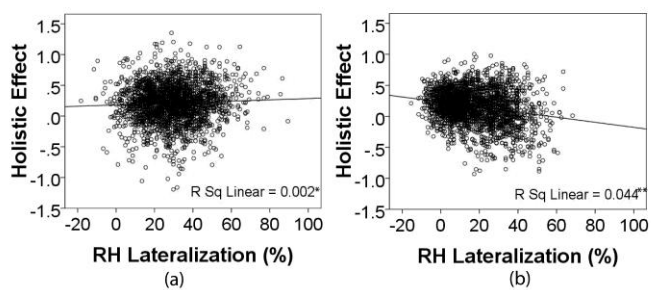

Figure 7: Relationship between RH lateralization and holistic processing in the (a) configural (b) featural tasks.

Thus, our data suggest that holistic processing (measured by the composite paradigm) and RH lateralization are separate processes that do not always go together. More specifically, the properties of the internal representation learned by the model can influence holistic processing and RH lateralization differently, depending on the task requirements.

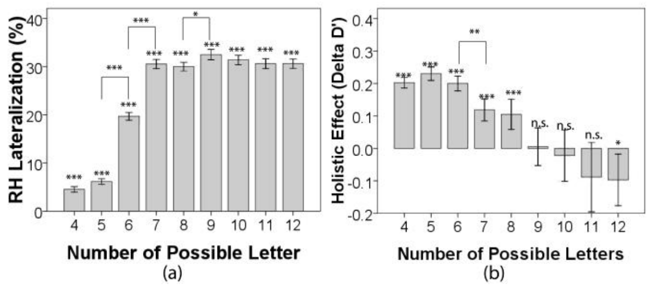

Figure 8: Featural tasks: (a) RH lateralization; (b) Holistic processing in the biased condition (comparisons with 0 and pair comparisons, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

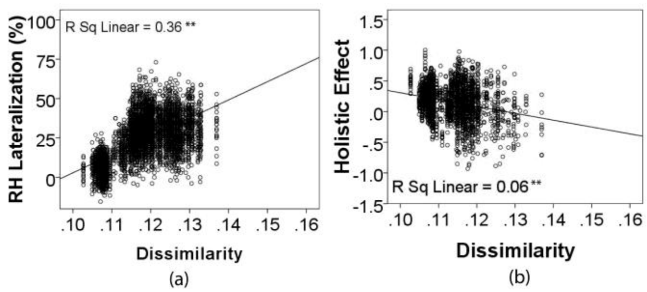

Figure 9: Featural tasks: Relationship between (a) stimulus dissimilarity and RH lateralization, and (b) stimulus dissimilarity and holistic processing.

Conclusion & Discussion

Holistic processing has long been considered as a property of RH processing (e.g., Bradshaw & Nettleton, 1981). Consistent with this view, it has been found that holistic face processing measured in the composite paradigm is linked to RH processing (e.g., Schiltz et al., 2010). However, recent research showed that Chinese character recognition expertise involves reduced holistic processing and increased RH lateralization (Hsiao & Cottrell, 2009), suggesting that holistic processing and RH lateralization may be separate processes that do not always go together.

Here we investigated the relationship between holistic processing and RH lateralization in configural and featural recognition tasks through computational modeling. Our model implements a theory of hemispheric asymmetry in perception, the DFF theory, which posits a LSF bias in the RH and a HSF bias in the LH; this model (or a variant) has been shown to be able to account for both RH lateralization and holistic processing in face recognition (e.g., Hsiao et al., 2008; Cottrell et al., 2002; Richler et al., 2007). Our data showed that although in both the featural and configural tasks, RH lateralization decreased with increasing stimulus similarity, in the featural tasks, holistic processing increased with increasing stimulus similarity, whereas no correlation between holistic processing and stimulus similarity was observed in the configural tasks. In addition, whereas RH lateralization and holistic processing were positively correlated in the configural tasks, in the featural tasks this correlation was negative. This effect suggests that the internal representation learned by the model in the recognition tasks has properties that can influence holistic processing and RH lateralization differently depending on the task requirements, for example, whether the task depends on featural or configural information.

This result has important implications for the research on visual cognition. For example, visual word recognition relies more on featural processing since configural information is not important for distinguishing words (e.g., Ge et al., 2006); consistent with our modeling data, recent studies showed that Chinese character recognition expertise involves RH lateralization and reduced holistic processing (Hsiao & Cottrell, 2009), whereas English word recognition expertise involves LH lateralization (e.g., McCandliss, Cohen, & Dehaene, 2003) and increased holistic processing (Wong et al., submitted) – a negative correlation between holistic processing and RH lateralization; this result is consistent with our data that in the featural recognition tasks there is a negative correlation between holistic processing and RH lateralization. In contrast, configural information has been shown to be more important for face recognition than the recognition of other types of objects (e.g., Farah et al., 1998), and thus holistic face processing has been found to be linked to RH lateralization (e.g., Schiltz et al., 2010), consistent with our data. Note however that both featural and configural information may be important for face recognition (e.g., Rotshtein et al., 2007); our modeling data suggest that the relationship between RH lateralization and holistic processing depends on the task requirements. Thus, Future work will examine the relationship between the two effects when both featural and configural information are important for recognition.

Note that holistic processing in visual cognition research has been measured in different ways; although the composite paradigm is the most common method, it has also been measured by, for example, the part-whole paradigm (Tanaka & Farah, 1993). Whether the effects observed here can also be observed in other paradigms requires further examinations. On a similar note, our result is not completely inconsistent with the holistic/analytic dichotomy proposal in the hemispheric asymmetry literature, as the definition of holistic processing can be broad to include concepts such as global, synthetic, or gestalt processing (e.g., Bradshaw & Nettleton, 1981). Nevertheless, our result suggests that a better description of RH processing may be needed.

In summary, in contrast to the well-accepted proposal that holistic processing is a property of RH processing, our modeling data suggest that holistic processing (measured by the composite paradigm) and RH lateralization are separate processes that do not always go together, depending on the task requirements.

Acknowledgement

We are grateful to the HKU Seed Funding Program for Basic Research (project #10400471 to J.H. Hsiao) and the

Research Grant Council of Hong Kong (project code: HKU 744509H and HKU 745210H to J.H. Hsiao). We thank Dr. Antoni B. Chan for helpful comments.

References

Bradshaw, J. L., Nettleton, N. C. (1981). "The nature of hemispheric specialization in man". *Behav. Brain Sci.*, *4*, 51–91.

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends Cogn. Sci.*, *10*, 159–166.

Burt, D. M. & Perrett, D. I. (1997). Perceptual asymmetries in judgments of facial attractiveness, age, gender, speech and expression. *Neuropsychologia*, *35*, 685-693.

Cheung, K. C. F., & Hsiao, J. H. (2010). Visual and Task characteristics may explain hemispheric asymmetry in visual word recognition. *Proc. of the Thirty-Second Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cottrell, G. W., Branson, K., & Calder, A. J. (2002). Do expression and identity need separate representations? *Proc. of the 24th Annual Cognitive Science Conference*.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two dimensional visual cortical filters. *J. Opt. Soc. Am. A*, 2, 1160-1169.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. N. (1998). What is “special” about face perception? *Psychol. Rev.*, *105*, 482–498.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition*, *103(2)*, 322-330.

Gauthier, I., & Tarr, M. J. (2002) Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *J. Exp. Psycho. Human*, *28*, 431-446.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training "Greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, *38*, 2401-2428.

Ge, L., Wang, Z., McCleery, J. P., & Lee, K. (2006). Activation of face expertise and the inversion effect. *Psychol. Sci.*, *17*, 12–16.

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

Harris, A., & Aguirre, G. K. (2008). The representation of parts and wholes in face-selective cortex. *Journal of Cognitive Neuroscience*, *20(5)*, 863-78.

Hsiao, J. H., & Cottrell, G. W. (2009). Not all expertise is holistic, but it may be leftist: The case of Chinese character recognition. *Psychol. Sci.*, *20(4)*, 455-463.

Hsiao, J. H., Shieh, D., & Cottrell, G. W. (2008). Convergence of the visual field split: hemispheric modeling of face and object recognition. *J. Cognitive Neurosci.*, 20(12), 2298-2307.

Hsiao, J. H., & Shillcock, R. (2006). Analysis of a Chinese phonetic compound database: Implications for orthographic processing. *J. Psycholinguist Res.*, *35*, 405-426.

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

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

Lades, M., Vorbruggen, J. C., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, R. P., & Konen, W. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE T. Comput.*, 42, 300-311.

Legendre, P., & Legendre, L. (1998). *Numerical ecology*. Amsterdam: Elsevier Science BV.

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

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends Cogn. Sci.*, *11*, 8–15.

Meadows, J. C. (1974). The anatomical basis of prosopagnosia. *J. Neurol. Neurosur. PS.*, *37*, 489 – 501.

Mondloch, C. J., Grand, R. L., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, *31*, 553–566.

Richler, J. J., Cheung, O. S., & Gauthier, I. (in press). Holistic processing predicts face recognition. *Psychological Science*.

Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2007). Distinguishing Between Perceptual and Decisional Sources of Holism in Face Processing. *Proc. of the 29th Annual Cognitive Science Conference*.

Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage*, *20*, 1609-1624.

Rotshtein, P., Geng, J. J, Driver, J., & Dolan, R. J. (2007). Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: behavioral and functional magnetic resonance imaging data. *J. Cogn. Neurosci*., *19(9)*, 1435-1452.

Sanger, T. (1989). An optimality principle for unsupervised learning. In Touretzky, D. (ed) *Adv. Neur. In.*, vol. 1, pp. 11-19, San Mateo: Morgan Kaufmann.

Schiltz, C., Dricot, L., Goebel, R., & Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *J. Vision*, *10(2):25*, 1–16.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Q. J. Exp. Psychol.*, 46A, 225–245.

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nat. Neurosci.*, *3*, 764-769.

Wong, A. C.-N., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Freenspon, E. (submitted). Holistic Processing of Words Modulated by Reading Experience.

Wong, A. C.-N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for face-like expertise with objects: Becoming a Ziggerin expert – but which type? *Psychol. Sci.*, *20(9)*, 1108-1117.