

國立臺灣大學理學院心理學研究所

博士論文

Graduate Institute of Psychology

College of Science

Nation Taiwan University

Doctoral Thesis

視覺系統對漢字的階層性處理

Hierarchical Processing of Visual Word Forms



高千惠

Chien-Hui Kao

指導教授：陳建中 博士，胡志偉 博士

Advisers: Chien-Chung Chen; Chih-Wei, Hue, Ph.D.

中華民國 100 年 01 月

January, 2011

國立台灣大學理學院心理學研究所

論文口試委員會審定書

高千惠 先生所提論文 The hierarchical processing of
visual word forms

經本委員會審議，符合 博 士學位標準，特此證明。

論文考試委員會

主席	<u>黃榮村</u>	
委員	<u>黃榮村</u>	<u>葉孝玲</u>
	<u>汪昱頌</u>	<u>胡志偉</u>
	<u>陳建中</u>	
	<u>姜自強</u>	
指導教授：	<u>陳建中</u>	<u>胡志偉</u>
所主任：	<u>翁饒禎</u>	

中華民國 100 年 01 月 12 日

Acknowledgement

「不積跬步，無以致千里；不積小流，無以成江海」《荀子》。

博士論文的完成象徵著一個階段的結束也是另一個階段的開始。在這個漫長的階段，首先要感謝我的兩位恩師，胡志偉老師與陳建中老師。從碩士班開始，便有幸跟隨胡志偉老師，從他身上學到的不僅僅只是專業，還有作為一個學者應有的態度與胸襟，而這樣的器宇是我最大的寶藏也是我在台大最大的收穫。胡老師總是在大處指引我正確的方向，而關於細節的部分卻給我最大的揮灑空間。在博士班的生涯，陳建中老師帶領我進入神經科學的領域，擴展了我的視野，恢弘了我的領域，在浩瀚的學海中帶領我走出一條可長可久的道路。有了這些基礎，我才有能力結合神經科學與最新的影像技術來完成這一點有限的研究。若我的研究真的有些微的貢獻，那多要歸功兩位老師的指導，若有任何的瑕疵，則應由我來負責。

而所有的口試委員們，包括：黃榮村校長、葉素玲老師、汪曼穎老師以及姜自強老師等提供我相當寶貴的意見，讓我的研究更加完備與豐富。他們思考與研究的深度和廣度，都是未來我所積極學習追隨的目標，在這裡表達我最深的謝意。

在博士班的生涯中，特別要感謝德佑學長，帶領我進入MRI的領域，耐心且親切地回答我任何相關的問題。另外，謝謝電機系的謝昭賢博士與MRI實驗室同仁，在這段時間提供最好的服務，確保MRI實驗室中各種儀器的穩定，這對我來說是極為關鍵。謝謝呂菁菁老師、Dr. Gabrenya、簡惠玲老師，總是給我溫暖的擁抱與鼓勵。謝謝身邊的朋友，明霓，你雖然在遠方，但我們的心是繫在一起的，

你對我的關心，並不因此而打了折扣。謝謝Chris，仔細幫我修改英文的句子。實驗室的好朋友們，榮瑜、琦琦總是給我歡樂，抒解我寫論文的疲憊。尹襄、冠銘、昱翔、羅明、盈炆、婉君、美靜、彥芳、佳儀等給我鼓勵與支持。還有高中的好姊妹們，尤其是淑如，你們為我的祈福與關心，是我前進的最大動力。

謝謝二姨媽、小阿姨跟三阿姨等長輩，因為你們的包容與支持，讓我無後顧之憂地向前邁進。謝謝阿本與狗狗嘟嘟，每週末陪著我上山下海，徜徉於大自然中，常常對我談古說今，這是我生命中最幸福的事情。最後感謝我的家人，尤其是我的父母與老弟，你們永遠在背後支持我，使我心無旁騖，可以盡情在浩瀚的學海中遨遊，我永遠以身為你們的家人為榮。

該感謝的親朋好友還很多，我無法一一列舉，就留下一個位置：□□□，就麻煩您自行填上，在此感謝你。



中文摘要

漢字是一種圖像文字，每個漢字都佔據固定的空間以做適當的安排，字形筆畫的結構是影響中文讀者在閱讀時的重要因素。本研究主要的目的是探討視覺系統對於漢字的處理歷程以及其神經機制。本研究以心理物理學的方式測量漢字的空間加乘功能、並採用功能性核磁共振造影的技術檢驗漢字在視覺區的表徵。在研究一中，我們操弄真字、假字、非字、甲骨文以及破碎字等刺激大小以及它們在視野的偏角。根據中文讀者對於這些漢字在偵測作業與區辨作業上的表現，我們使用空間加乘模型推估漢字在視覺系統中的機制。實驗結果顯示，在偵測作業中，參與者對於文字大小的偵測閾值不受文字類別影響，參與者對目標字的閾值會隨著刺激的變小而提高，且斜率接近-1/2。在區辨作業中，參與者對於目標字的閾值也是隨著刺激的變小而提高，且斜率接近-1，且參與者在區辨真字與假字或是非字的閾值高於區辨真字與甲骨文或是破碎字。再者，隨著視野偏角的增大，參與者對於文字的偵測與區辨力都會變差。這個結果顯示在偵測作業中，視覺系統主要受到漢字之細部特徵的影響；在區辨作業中，視覺系統則則是會受到漢字之熟悉特徵的影響。在研究二中，我們採取網膜拓樸對應派典找出視覺區相對應於漢字處理的特定表徵。我們比較大腦視覺區在"空白背景的旋轉楔形與擴大環狀之漢字"與"破碎字背景的旋轉楔形與擴大環狀之漢字"的激發狀況。實驗結果顯示，"空白背景的旋轉楔形與擴大環狀之漢字"會激發初級視覺皮質區與視覺顳葉區，這兩個區域都顯示了視覺區對漢字網膜拓樸的特性。"破碎字背景的旋轉楔形與擴大環狀之漢字"則只會激發視覺顳葉區，尤其是梭狀回對整字有網膜拓樸的特性；亦即梭狀回對於漢字在視網膜的位置具有特定的對應性。再者，相較於破碎字，梭狀回偏好那些呈現在視野中央的文字。在研究三與四中，我們以心理物理測量方法與功能性核磁共振造影比較參與者對於正立漢字與倒立漢字的表現。在行為結果上，我們發現參與者對於倒立漢字的配對表現顯著地比正

立漢字差；而此我們稱之為倒立效果。我們同樣也在漢字部件上觀察到倒立效果，但在非字或是甲骨文上卻沒有這樣的倒立效果。這個結果顯示視覺系統會分析整字與部件的空間編碼訊息。配合神經造影的研究結果顯示，視覺系統是以階層的方式處理漢字，右腦的視覺頂葉區主要分析漢字的細部特徵，如，筆畫；右腦的視覺顯葉區處理漢字部件之間的空間關係；左腦的梭狀回則是處理漢字整體的空間訊息。總結，從視覺頂葉區到腹側視覺區，視覺系統會以階層的方式處理漢字。

關鍵字： 漢字、空間編碼、倒立效果、視覺皮質區、心理物理、功能性核磁共振造影。



Hierarchical Processing of Visual Word Forms

Chien-Hui Kao

Abstract

The present study is to investigate the processing of Chinese character in the visual system and its neural network. In study 1, we used a dual task paradigm with both detection and discrimination tasks in spatial summation to measure the spatial constraints for visual word form perception. Five types of stimuli (real, pseudo-, non-, *Jiagu* and scrambled characters) were used in the experiments. The detection thresholds for the same stimulus size and the same eccentricity were the same for all types of stimuli. When the target size was small, the detection threshold of a character decreased with the increase in its size, with a slope of $-1/2$ on log-log coordinates, up to a critical size at all eccentricities and for all stimulus types. Beyond the critical size, there was little, if any, improvement of visual word form detectability as target size further increased. That is, the detectability is based on local feature analysis regardless of character types. The contrast thresholds for discriminating real and pseudo- or non-characters were higher than for discriminating real characters and *Jiagu* characters or scrambled characters. The results suggested that the discrimination threshold is mediated by well-learned forms of components in a character. In study 2, we used a 2 (“rotating wedge of characters” vs. “expanding ring of characters”) by 2 (“with a blank background” vs. “with a scrambled background”) design with fMRI to define the visual cortical regions produce unique representations for Chinese characters. The ventral occipital region showed a retinotopic representation for character stimuli as we found

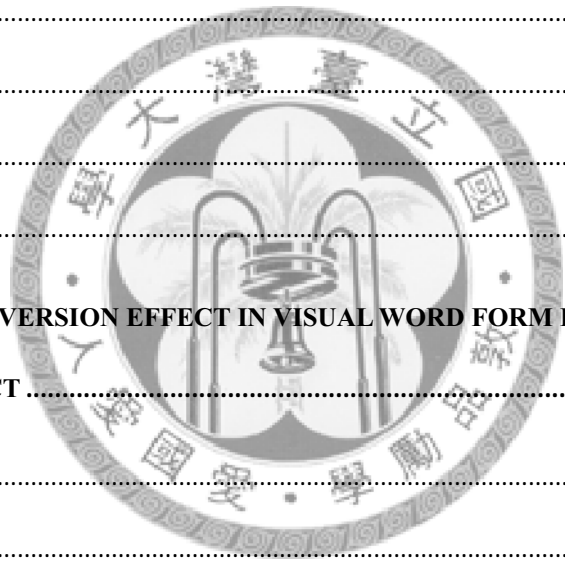
the fusiform gyrus is selective for the contralateral visual fields and has a centre-to-periphery map with a preference for characters presented in fovea. This implies that the fusiform gyrus is selectively to the spatial location information of stimuli. In study 3 and 4, we examined the spatial configuration processing in the visual system for a character by comparing the visual performance for upright character with that for their inverted versions with psychophysics and fMRI. We showed a robust inversion effect for real character regardless of eccentricities. Such an inversion effect was absent for non-characters. The inversion effect was also found in radicals or components in a character. These results suggested that a visual analysis of characters and their components involved a spatial configuration processing. Our neuroimaging evidence showed that the left fusiform gyrus analyzes the spatial configuration of components in a character while the right occipitotemporal regions analyze the spatial configuration of strokes in a component. In addition, the right occipitoparietal region analyzes local features of a character. Together, we demonstrated a hierarchical processing of visual word forms in the visual cortex.

Keywords: Chinese character, spatial configuration, inversion effect, visual cortex, psychophysics, fMRI.

Table of Contents

ACKNOWLEDGEMENT.....	I
中文摘要.....	III
ABSTRACT	V
TABLE OF CONTENTS.....	VII
LIST OF TABLES.....	XI
LIST OF FIGURES	XIII
CHAPTER 1 : INTRODUCTION.....	1
1.1 VISUAL PROCESSING FOR VISUAL WORD FORMS.....	2
1.2 HIERARCHICAL PROCESSING IN THE VISUAL CORTEX.....	5
1.2.1 Hierarchical processing in the visual cortex.....	5
1.2.2 Functional segregation in the ventral cortex.....	7
1.3 VISUAL WORD FORM PERCEPTION IN THE VISUAL CORTEX.....	8
1.3.1 The processing of alphabetic language in the visual cortex.....	8
1.3.2 The processing of orthographic language in the visual cortex.....	9
1.4 OVERVIEW OF THIS THESIS.....	11
CHAPTER 2 : SPATIAL SUMMATION OF VISUAL WORD FORM PROCESSING.....	15
2.1 METHODS.....	19
2.2 RESULTS.....	22
2.3 DISCUSSION.....	35
2.4 SUMMARY.....	44

CHAPTER 3 : RETINOTOPIC MAPPING OF VISUAL WORD FORMS IN THE VISUAL	
CORTEX	45
3.1 METHODS	47
3.2 RESULTS	53
3.3 DISCUSSION	64
3.4 SUMMARY	67
CHAPTER 4 : THE INVERSION EFFECT IN VISUAL WORD FORM PROCESSING: THE	
SPATIAL CONFIGURATION	69
4.1 METHODS	72
4.2 RESULTS	78
4.3 DISCUSSION	93
4.4 SUMMARY	103
CHAPTER 5 : THE INVERSION EFFECT IN VISUAL WORD FORM PROCESSING: THE	
COMPONENT EFFECT	105
5.1 METHODS	108
5.2 RESULTS	116
5.3 DISCUSSION	127
5.4 SUMMARY	133
CHAPTER 6 : GENERAL DISCUSSIONS	135
REFERENCES	143
APPENDICES.....	163
APPENDIX 1: CHARACTER STIMULI IN CHAPTER 2	163
APPENDIX 2: RETINOTOPIC ACTIVATION FOR OTHER 5 OBSERVERS IN CHAPTER 3	166



APPENDIX 3: CHARACTER STIMULI IN CHAPTER 4 AND 5.....178

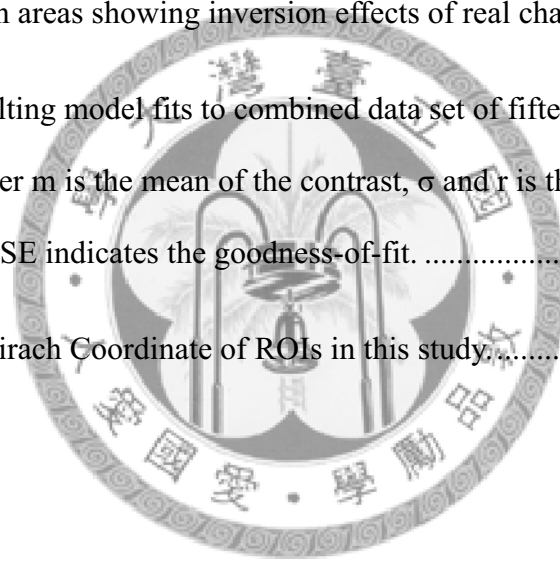
CURRICULUM VITAE.....181





List of Tables

Table 2.1	RMSE and SSE for each type of stimuli.....	36
Table 4.1	The resulting model fits to combined data set of eleven observers. The parameter μ is the mean of the contrast, σ and ρ is the scale parameter, and RMSE indicates the goodness-of-fit.	83
Table 4.2	The brain areas showing differential between characters and scrambles .	89
Table 4.3	The brain areas showing inversion effects of real characters	93
Table 5.1	The resulting model fits to combined data set of fifteen observers. The parameter m is the mean of the contrast, σ and r is the scale parameter, and RMSE indicates the goodness-of-fit.	117
Table 5.2	The Talairach Coordinate of ROIs in this study.....	121





List of Figures

- Figure 2.1 Examples of the stimuli used in the study. a) real character; b) pseudo-character; c) non-character; d) *Jiagu* character; e) scrambles.....21
- Figure 2.2 The contrast threshold for all types of character-like stimuli at different eccentricities in the detection task. Each panel represents different eccentricities in the detection task. Each row denotes one observer and each column denotes each eccentricity. The spatial summation curves are fits of the model defined in the visual word form. The error bars are one standard error of the means.....26
- Figure 2.3 The contrast threshold for all types of character-like stimuli at different eccentricities in the discrimination task. Each column denotes one observer and each row denotes each eccentricity. The spatial summation curves are fits of the model defined in the visual word form. The error bars are one standard error of the means.28
- Figure 2.4 The detectability for different eccentricities. Each row denotes one observer and each column denotes each stimuli type. From top to bottom is real characters, non-characters, *Jiagu* characters and scrambles, respectively. The spatial summation curves are fits of the model defined in the visual word form. The error bars are one standard error of the means.32
- Figure 2.5 The discriminability for different eccentricities. Each row denotes one observer and each column denotes each stimuli type. The spatial

summation curves are fits of the model defined in the visual word form.
The error bars are one standard error of the means.34

Figure 2.6 ω_{T2} across three observers in the detection task. The error bars are one
standard error of the means.....40

Figure 2.7 ω_{T1} and ω_{T2} across three observers in the discrimination task. The error
bars are one standard error of the means.41

Figure 3.1 The retinotopic paradigm of Chinese characters. A) the rotating wedge of
characters with a blank background; B) the rotating wedge of characters
with a scrambled background; C) the expanding ring of characters with a
blank background; D) the expanding ring of characters with a scrambled
background.....49

Figure 3.2 Flatmaps centered near the occipital pole of the left hemisphere (LH) and
the right hemisphere (RH) for observer AEW. The retinotopic areas
V1-V3 (red, green and blue outlines) was defined by the rotating-wedge
and expanding-ring checkerboards. The gyral and sulcal landmarks are
shown as light-gray and dark-gray shading, respectively.52

Figure 3.3 Retinotopic activations for the A) rotating-wedge and B) expanding-ring
checkerboards for three typical observers. Radial color map segregates
the left field(LF) and the right field(RF). Concentric color map separates
the fovea and periphery. The outlines of red, green and blue denote V1,
V2, and V3, respectively. LH (left hemisphere), RH (right hemisphere) 56

Figure 3.4 Retinotopic activations for A) rotating-wedge and B) expanding-ring
characters with a blank background for three typical observers. Radial

color map segregates the LF and the RF. Concentric color map separates the fovea and periphery. V1 to V3 are defined by the rotating-wedge and expanding-ring checkerboards. The magenta, yellow, black and white borders denote the LO, FG, VWFA, and FFA, respectively. LH (left hemisphere), RH (right hemisphere).....60

Figure 3.5 Retinotopic activations for A) rotating-wedge and B) expanding-ring characters with a scrambled background. Radial color map segregates the LF and the RF. Concentric color map separates the fovea and periphery. V1 to V3 are defined by the rotating-wedge and expanding-ring checkerboards. The magenta, yellow, black and white borders denote the LO, FG, VWFA, and FFA, respectively. LH (left hemisphere), RH (right hemisphere).....63

Figure 4.1 Examples of the stimulus types used. a) an upright real character; b) an upright non-character; c) an inverted real character; d) an inverted non-character..... 74

Figure 4.2 The Gaussian accumulative distribution functions as a fitting model for the data. Proportional correct responses of upright real characters, inverted real characters, upright non-characters, and inverted non-characters at different contrast thresholds. a) real character stimuli were presented at 1° eccentricity, b) real characters were presented at 4° eccentricity, c) non-characters were presented at 1° eccentricity, d) non-characters were presented at 4° eccentricity. In each panel, the solid denotes the proportion of correct responses across eleven observers' data. * $p < .05$ and ** $p < .01$. Error bars are ± 1 SEM.....81

Figure 4.3 Proportional correct responses of upright real and non-characters at different contrast thresholds. a) real characters and non-characters were presented at a 1° eccentricity, b) real characters and non-characters were presented at a 4° eccentricity. * $p < .05$ and ** $p < .01$82

Figure 4.4 (a) Activation maps are shown in the standardize brain slices for the real characters versus their scrambled versions. Orange color denotes the average activation pattern. (b) The real characters activation depicted in flatmaps centered on the occipital poles. (c) Activation maps are shown in the standardized brain slices for the non-characters versus their scrambled versions. (d) The non-characters activation depicted in flatmaps centered on the occipital poles. (e) Differential activation between real-characters and non-characters in flatmaps. The blue border denotes the extent of the activated areas in the FCA. The cyan border denotes the extent of the activated areas in the LOCA. The magenta border denotes the extent of the activated areas in the OPCA. Gyral (light) and sulcal (dark) denoted by shading. Yellow-orange coloration codes the negative log base 10 of the significance value of a 2-side t -test (threshold at $p < .05$). L: left hemisphere; R: right hemisphere.88

Figure 4.5 Percentage change of activation between the upright and inverted real and the non-characters in FCA and LOCA. The error bars denote the standard error calculated across participants.92

Figure 4.6 (a) Differential activation between upright and inverted real characters and (b) differential activation between upright and inverted non-characters depicted in flatmaps centered on the occipital poles. Blue

borders: FCA; cyan borders: LOCA; magenta borders: OPCA. Activation colors as in Figure 4.4.92

Figure 4.7 Activation maps for the upright versus inverted real characters on the occipital-poles flatmaps. The Talairach coordinates of each study were presented. The black border denotes the differential activated region between real- and non-characters. Cyan borders: LOCA; magenta borders: OPCA. Colored symbols indicated mean locations of activations in other studies (see key). Activation colors as in Figure 4.4.99

Figure 5.1 An example of multivariate analysis procedure in this study. A) The VWFA localizer for observer WSY. A voxel was considered to have a significant activation if its *t*-test has a p-value smaller than the α -level of .005. These voxels were used as feature parameters and were included in the classification analysis. B) The observers viewed all character types of stimuli. We extracted feature parameters for classification analysis. C) The classifier we used is a 2-dimensional feature space of voxel patterns. Each dot represents a voxel pattern and the color of a dot indicates stimulus type. D) Pattern vectors that were evoked by the presentation of stimuli across selected voxels from three out of four runs were assigned to a training data set. Each brain pattern vector was labeled according to its experimental condition. E) Pattern vectors from training data set were used to train a linear support vector classifier that maps between brain activity and experimental conditions. We tested the classifier on activity patterns evoked by the same character types from the test data set. F) The remaining 4th run was assigned to a test data set. Four-fold

cross-validation conducted by repeating this procedure in which each run was assigned to the test data set. The accuracy was averaged over these four iterations. 115

Figure 5.2 The Gaussian accumulative distribution functions as a fitting model for the data. Proportional correct responses of upright and inverted real, non-, lexical, non-lexical, and *Jiagu* characters at different contrast thresholds for 11 observers. a) real characters, b) non-characters, c) lexical components, d) non-lexical components, e) *Jiagu* characters. In each panel, the solid denotes the proportion of correct responses across all observers' data. * $p < .05$ and ** $p < .01$. Error bars are the SEM across all observers. 119

Figure 5.3 The regions of interest for the observer WSY. L: left hemisphere; R: right hemisphere. 120

Figure 5.4 Classification performance for character types versus scrambled lines in different regions of interest. A) Prediction accuracy for the discrimination of real characters versus scrambled characters. B) Prediction accuracy for the discrimination of non-characters versus scrambled characters. C) Prediction accuracy for the discrimination of lexical components versus scrambled characters. D) Prediction accuracy for the discrimination of non-lexical characters versus scrambled characters. Error bars are the SEM across all observers. LH: left hemisphere; RH: right hemisphere. * $p < .05$ and ** $p < .01$ 123

Figure 5.5 Prediction accuracy for the discrimination of real characters versus character components in different regions of interest. A) Prediction

accuracy for the discrimination of real characters versus lexical components. B) Prediction accuracy for the discrimination of real characters versus non-lexical components. Error bars are the SEM across all observers. LH: left hemisphere; RH: right hemisphere. * $p < .05$ and ** $p < .01$ 124

Figure 5.6 Prediction accuracy for the discrimination of upright character types versus inverted counterparts in different regions of interest. A) Prediction accuracy for the discrimination of upright versus inverted real characters. B) Prediction accuracy for the discrimination of upright versus inverted non- characters. C) Prediction accuracy for the discrimination of upright versus inverted lexical components. D) Prediction accuracy for the discrimination of upright versus inverted non-lexical components. Error bars are the SEM across all observers. LH: left hemisphere; RH: right hemisphere. * $p < .05$ and ** $p < .01$ 126

Figure 6.1 Hierarchical processing model of visual word form processing. EVA: early visual areas; OPCA: occipitoparietal character area; LO: lateral occipital region; FG: fusiform gyrus..... 139

Chapter 1 : Introduction

Visual word form is one type of visual objects that we have been exposed to extensively since childhood. It is common for a person in modern society to process thousands of words/characters every day. With such amount of practice, it is not surprised that a fluent reader can recognize words with remarkable speed and accuracy (Grill-Spector & Kanwisher, 2005; Rayner & Pollatsek, 1989). While there have been many behavioral and neural brain imaging studies on visual word recognition, scientists have little knowledge about how our visual system achieves such efficiency and accuracy in visual word form processing (Pelli, Burns, Farell, & Moore-Page, 2006).

The purpose of this study is to investigate the visual word form processing in the visual system with psychophysics and function magnetic resonance imaging (fMRI) methods. We intend to achieve the following goals:

- 1) To demonstrate that there is a specific mechanism underling visual word perception.
- 2) To determine the neural mechanism underlying the visual word processing with fMRI.
- 3) To clarify the role of the local feature analysis and the global structure analysis in visual word form processing.

To achieve these goals, in Chapter 1, we first review some behavioral studies on visual word form perception and provide three hypotheses (session 1.2). In order to understand how the human brain processes visual word form, we review the physical properties of the visual cortex and its functional modalities for processing objects (session 1.3). In session 1.4, we review the neuroimaging studies that have examined several visual cortical regions which may play an important role in visual word processing. In reviewing previous studies we aim to accomplish an integrative understanding of the relationship between perceptual behaviors and neural activities for visual word processing.

In Chapters 2 to 5, we present the results of our experiments. In Chapter 2, we apply the spatial summation to determine the factors that affect visual word form perception near the perceptual threshold with psychophysics. In Chapter 3, we use the retinotopic paradigm to examine how the visual cortical regions produce unique representations for Chinese characters. In Chapters 4 and 5, we examine the role of spatial configuration in processing Chinese characters with psychophysics and fMRI. In Chapter 6, we summarize the results of these experiments and provide a hierarchical processing model to account for visual word forms in the visual system.

1.1 Visual processing for visual word forms

One of the main issues in the visual word form perception literature is whether there is a specialized mechanism for visual word form analysis in the visual system and, if so, what their response properties are (Cattell, 1886; Pelli, et al., 2006). Three hypotheses of visual word form perception have been proposed in the literatures. The first, the local feature analysis hypothesis, suggests that the recognition of a visual word

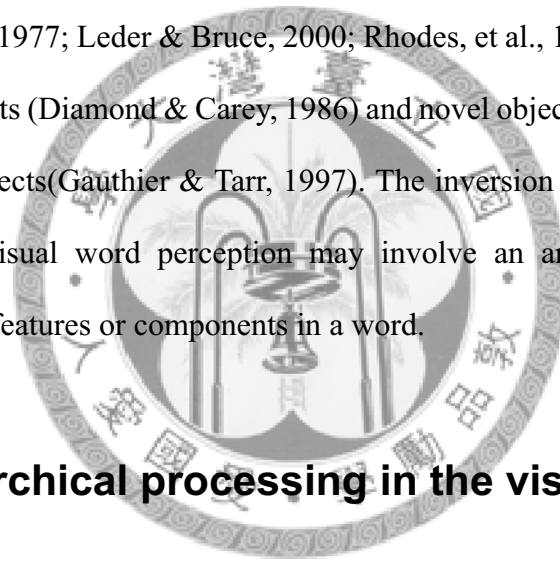
is based on the analysis of individual letters, which in turn is relied on the analysis of local image features such as the length or the width of the strokes (Grill-Spector & Kanwisher, 2005; Martelli, Majaj, & Pelli, 2005; Pelli, Farell, & Moore, 2003). That is, a character is nothing more than a set of strokes for the visual system. Several neural network models, such as the parallel distributed processing (PDP, Rumelhart & McClelland, 1986), share a similar view. Pelli et al. (2006) compared the performance of a human observer to that of an ideal observer (e.g., the efficiency) in detecting a letter written in different fonts (e.g., Helvetica, Sloan, and Kunstler) or from different languages (e.g., Arabic and Chinese) embedded in visual noise. They found that the detection and identification efficiency of a human observer decreased with the increase of the complexity (the ratio of the “ink” area in a perimeter of the stimulus) of stimuli. A number of factors, such as viewing conditions and languages had no effect on the detection and identification efficiency. This result implies that letter detection and identification is limited only by the local image features. Pelli et al. (2003) also found that the word identification efficiency was inversely proportional to word length and was independent of how many possible words the test word is drawn from. Hence, the word identification performance can be easily accounted for by a model of local feature analysis.

The global shape template hypothesis, on the other hand, assumes that the shape of the whole word is critical for visual word recognition (Healy & Cunningham, 1992; Perea & Rosa, 2002). That is, the visual system directly analyzes the spatial relationship of strokes relative to the whole character. Perea and Rosa (2002) found that alternating a visual word’s size (e.g., LeTtEr) increased participants’ reaction time

in the lexical decision task. Thus they concluded that the global word shape plays an important role in lexical access.

Finally, the spatial configuration analysis hypothesis assumes that, while local feature analysis is necessary at the first step of visual word form processing, the visual word recognition requires an analysis of spatial relationship among features in a word (Pammer, Lavis, Cooper, Hansen, & Cornelissen, 2005). With other types of familiar objects, such as faces, the visual system tends to analyze the spatial configuration, or spatial relationship among image elements, rather than the image elements themselves (Carey & Diamond, 1977; Pammer, et al., 2005; Tanaka & Farah, 1993). Pammer et al. (2005) showed a string of symbols to their observers and then asked their observers to identify which one of the two subsequently presented strings had been shown previously. They reported that the performance in this symbol-string task was highly correlated with that in a lexical decision task in which the observer was asked to determine which one of the two strings was a word. Since there was no semantic information needed in the symbol-string task, this high correlation suggested that the ability to analyze spatial relation in the symbol-string task was also essential in visual word processing. In orthographic words, Yeh and Li (2002) found that fluent Chinese readers tended to judge visual similarity among characters with the spatial configuration of a character. The global structure and components of a character could affect the participants' efficiency in the visual search. Yeh, Li, Takeuchi, Sun, and Liu (2003) showed that skilled readers of characters (e.g., Taiwanese or Japanese college students) tended to sort characters with similarities in global spatial relationships among character components, while no Chinese readers (e.g., American college students) tended to sort characters with similarities in character strokes or components.

Recently, researchers found that the matching performance was better for upright characters than for inverted characters (Wang & Kuo, 2007). Since the elements in the upright images are the same as in the inverted images and only the spatial relationships among image elements are destroyed by the inversion, such impairment of performance indicates that the function of the spatial configurations is disrupted in processing inverted images (Leder & Bruce, 2000; Rhodes, Brake, & Atkinson, 1993; Yin, 1969; see review by Rossion & Gauthier, 2002). This effect was called the inversion effect, and it has been used as evidence for supporting the role of spatial analysis of an object. This inversion effect has been found in recognizing human faces in normal populations (Carey & Diamond, 1977; Leder & Bruce, 2000; Rhodes, et al., 1993; Yin, 1969), dogs in dog training experts (Diamond & Carey, 1986) and novel objects in observers trained to identify those objects (Gauthier & Tarr, 1997). The inversion effect in visual words suggests that the visual word perception may involve an analysis of the spatial relationship among features or components in a word.



1.2 Hierarchical processing in the visual cortex

1.2.1 Hierarchical processing in the visual cortex

Hierarchical processing (Riesenhuber & Poggio, 1999, 2002) is a convenient way of understanding how the visual system derives the perception of objects from simple features. The hierarchical processing model assumes that there are several processing stages in the visual system. Neurons in each stage combine responses from neurons in the preceding stages. As a result, each neuron tunes to more complex features than those in the previous stages. One classic example of hierarchical processing is from the

seminal study by Hubel and Wiesel (1962, 1968). Hubel and Wiesel (1962, 1968) used single cell recording techniques to explore the properties of neurons in V1. They proposed that in the primary visual cortex, complex cells are constructed by convergent projections from simple cells, while simple cells are constructed by the convergent projections from the LGN. The results demonstrated a hierarchical processing from LGN to the cells in the primary visual cortex.

Further downstream, it is known that the occipitotemporal region responded more strongly when participants viewed images of complex objects, such as faces, chairs, and houses, than when they viewed textures, or scrambled objects (Grill-Spector et al., 1998; Hanson, Matsuka, & Haxby, 2004; James, Culham, Humphrey, Milner, & Goodale, 2003; Malach et al., 1995; Ostwald, Lam, Li, & Kourtzi, 2008). For example, Ostwald et al. (2008) showed that activations in the early visual regions were influenced by local features, such as dipole orientation, of a Glass pattern (Glass, 1969), while those in the lateral occipitotemporal regions were determined by the global forms of an object. Similarly, Kourtzi and Huberle (2005) also found the early visual areas are sensitive to the local orientations of contour elements, whereas the occipitotemporal region is sensitive to the global form of the contours. Grill-Spector et al. (1998) showed that scrambled images increased the activation in the early visual areas but decreased the activation in the higher visual areas. That is, V1 is sensitive to the local features of an object and higher visual areas are sensitive to the global shape of an object. Furthermore, neuroimaging studies also showed visual word forms are processed in a hierarchical way in the left occipitotemporal regions (Dehaene, Cohen, Sigman, & Vinckier, 2005; Dehaene et al., 2004; Vinckier et al., 2007). Neurons in the occipitotemporal regions are selective to oriented bars, to letters, bigrams, and finally

quadrants from the posterior to anterior regions. These electrophysiological and brain imaging studies of the visual system provide strong evidence that support Marr's hierarchical model (Kontsevich & Tyler, 1999).

To sum up, the hierarchical processing suggests the following dimensions and sequence of visual processes: (1) the visual system processes information in several stages; (2) each stage of the visual processing tunes in to specific visual features; (3) each stage of the visual processing integrates information from the preceding stage; and (4) in turn tunes in to more complicated features than the preceding stage.

1.2.2 Functional segregation in the ventral cortex

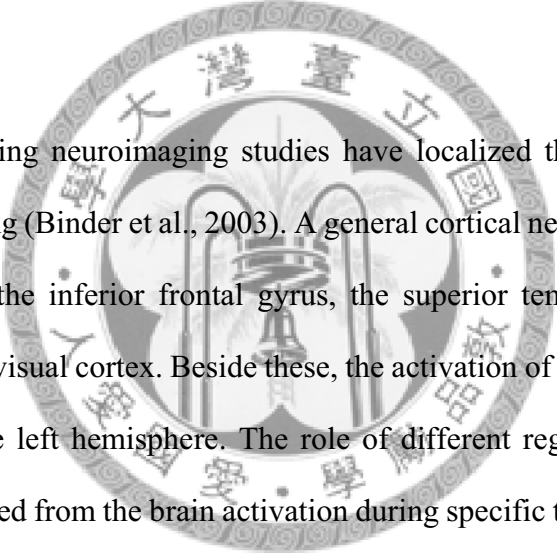
The visual areas in the ventral cortex show the highly functional segregation for object processing. For instance, the posterior fusiform gyrus responds to faces better than to other objects and is often called the fusiform face area (FFA, Kanwisher, McDermott, & Chun, 1997) in the literature. The middle fusiform gyrus is sensitive to linguistic material and is known as the visual word form area (VWFA, L. Cohen et al., 2000). The parahippocampus is sensitive to layout of space and is known as the parahippocampus place area (PPA, Epstein & Kanwisher, 1998). Furthermore, Mechelli, Price, Noppeney, and Friston (2003) used a dynamic causal model to investigate the functional selection of object perception in the ventral cortical regions. They found that the connectivity from inferior occipital cortex to the "house-responsive" areas was stronger during the perception of houses than other categories. Similarly, the connectivity from the inferior occipital cortex to the "face-responsive" areas and the "chair-responsive" areas was stronger during the perception of faces and chairs, respectively. Evidence from lesion studies also

suggests that damage in the occipitotemporal regions results in a variety of recognition deficits (Farah, 1990). These include object agnosia (the inability to recognize objects), prosopagnosia (the inability to recognize faces), and alexia (the inability to recognize visual words) (Goodale & Milner, 1992).

1.3 Visual word form perception in the visual cortex

1.3.1 The processing of alphabetic language in the visual

cortex



A rapidly growing neuroimaging studies have localized the neural circuits for visual word processing (Binder et al., 2003). A general cortical network for visual word processing includes the inferior frontal gyrus, the superior temporal gyrus and the ventral stream in the visual cortex. Beside these, the activation of alphabetic words was mostly located in the left hemisphere. The role of different regions for visual word processing was inferred from the brain activation during specific tasks. It was proposed that visual word processing is processed from the occipitotemporal regions to the inferior frontal gyrus. Researchers argued that the left occipitotemporal region is the first stage of the processing of visual word forms (Vinckier, et al., 2007). More specifically, many studies found that the left fusiform gyrus is robustly activated in the visual word recognition tasks (e.g., Beauregard et al., 1997; L. Cohen, et al., 2000; L. Cohen et al., 2002; Dehaene, Le Clec, Poline, Le Bihan, & Cohen, 2002; Polk & Farah, 2002). Cohen and his colleagues (2000, 2002) found that visually presented words produced a greater activation in the left fusiform gyrus than nonsensical letter strings

(non-words) or checkerboard patterns did. Further studies found that the left fusiform gyrus was activated by visually presented words irrespective of their typographical case, presentation location, color and size, but not by aurally presented words (Dehaene, et al., 2002; Polk & Farah, 2002). Hence, Cohen et al. (2000) argued that this region is specialized for the abstract word representations, and referred to it as the “visual word form area” (VWFA). However, there are some disagreements on such an interpretation (e.g. (Jobard, Crivello, & Tzourio-Mazoyer, 2003; Price & Devlin, 2003). For example, Jobard et al. (2003) furthered the notion of the general processing area and postulated that the VWFA is responsible for segmenting a visual stimulus into sub-units by detecting familiar forms embedded in the stimulus, such as familiar combinations of letters and figures. The role of the VWFA is still under debate.

1.3.2 The processing of orthographic language in the visual cortex

Unlike alphabetic languages in which an alphabetic word is constructed by 26 letters, a Chinese character is represented by a logographic script. In modern Chinese, most of the words are of two syllables, and except in rare cases, a character represents a syllable. That is, a character can be viewed as a building block or morpheme and directly represents its phonology and meaning. Over 90% of characters are compounds, comprising two or more constituents called components (or radicals). There are approximately 20 distinctive strokes and 200 components in the Chinese writing system (S. Y. Wang, 1973). A character is composed of several radicals or components that are composed of different strokes or stroke patterns. It has been estimated that a

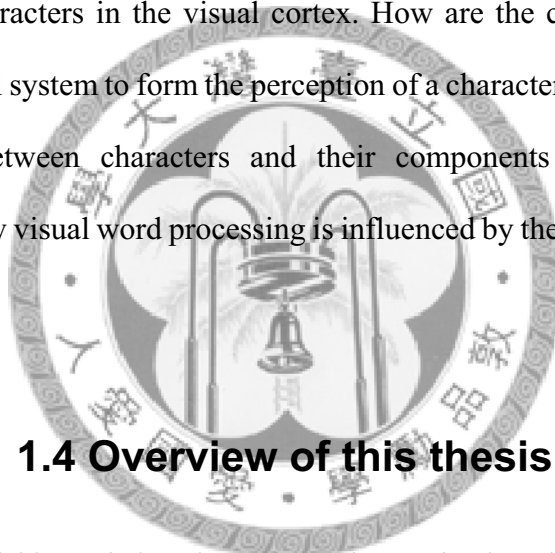
college student knows about 5,100 characters (C.W. Hue, 2003). With such complex construction characters and a vast vocabulary, for a Chinese language learner, the mastery of written characters is the most difficult task (Everson, 1998).

Despite the different written structures between alphabetic languages and orthographic languages, many neuroimaging studies have shown both alphabetic and orthographic languages share common neural circuits (Tan, Laird, Li, & Fox, 2005). In addition to these common regions, it has been found that orthographic languages, such as Chinese characters or Japanese Kanji, also activate the bilateral dorsal stream and occipitotemporal regions (Dong et al., 2000; Ino, Nakai, Azuma, Kimura, & Fukuyama, 2009; Kuo et al., 2003; Tan, Feng, Fox, & Gao, 2001; Tan et al., 2001). Therefore, the bilateral visual cortical regions may involve the spatial analysis of orthographic characters. While these studies showed the extensive activation of visual cortical regions, they failed to analyze or separate the functions of different loci of activation in the visual cortex during visual word form processing. The mechanisms of visual processing in character perception were not clear.

More recently, we presented a study that explored the functions of the visual cortex when processing visual word form and found that visual word perception is carried out in a hierarchical way (Kao, Chen, & Chen, 2010, also see chapter 4). We measured brain activation when viewing Chinese characters and non-characters and their inverted versions, using fMRI. The left fusiform gyrus showed a significant differential activation between upright and inverted real characters. This region also showed a differential activation between upright real and non-characters. This suggests that this region is sensitive to the spatial configuration of a character. The lateral occipital region showed a differential activation between upright and inverted real

characters, but not for the difference between upright real and non-characters, suggesting that this region is responsible for the local components rather than the character themselves. The occipitoparietal regions showed character selective activation when compared with scrambled lines. This region may be responsible for local features in a character. This is the first study to reveal the different functions of the visual regions when processing Chinese character information. It implies that the visual system may analyze words in a hierarchical way.

Nevertheless, our prior study did not address the relationship between components and characters in the visual cortex. How are the components combined together by the visual system to form the perception of a character? The investigation of the relationships between characters and their components would enhance our understanding of how visual word processing is influenced by the components and their interplay.



1.4 Overview of this thesis

The first goal of this study is to investigate the mechanism that underlies the visual word perception. More specifically, it aims to determine whether processing mechanisms exist for Chinese character perception and whether the perception of Chinese characters is based on the local feature analysis, the global shape template or the spatial configuration of a word.

In Chapter 2, we used a spatial summation paradigm to explore the mechanisms for character perception. The spatial summation paradigm has been used to test of the mechanisms of face perception (Tyler & Chen, 2006). Here we investigate character

legibility as a function of size, contrast, configuration and eccentricity in the detection and the discrimination tasks. Thus, the factors that affect the visual information available for Chinese character perception are still entirely unexplored. We attempt to fill the vacancy by modeling the spatial properties of visual word form processors in the visual system.

Two related questions emerge at this point: whether representations of character perception evoke distinct responses in the visual system? How the visual cortical areas produce unique representations for Chinese characters?

In Chapter 3, we apply a retinotopic mapping paradigm to identify the mechanism of hierarchical processing of character perception in the ventral occipital regions. A 2 (“rotating wedge of characters” vs. “expanding ring of characters”) by 2 (“with a blank background” vs. “with a scrambled background”) design was involved in this study. There are two reasons to use retinotopic mapping methodology in this study. First, previous studies have found that higher visual regions have certain retinotopic properties for object perception (Grill-Spector, et al., 1998; Hasson, Avidan, Deouell, Bentin, & Malach, 2003; Hasson, Levy, Behrmann, Hendler, & Malach, 2002; Levy, Hasson, Avidan, Hendler, & Malach, 2001). For example, FFA that responded to faces showed a foveal field bias and PPA that responded more strongly to houses and scenes showed a peripheral field bias (Hasson et al., 2002, 2003). Second, the retinotopic mapping identifies stimulus-driven areas in the visual cortex (Brewer, Liu, Wade, & Wandell, 2005; Wandell, Dumoulin, & Brewer, 2007). The effect of character perception in the ventral cortex can be attributed to the visual information in characters, but not to the phonological and semantic information.

A further set of questions leads to our third and fourth studies. These questions are as follows: How does our visual system process the spatial configuration of a character? Do the ways in which the features and components of a character are combined depend on the spatial configuration among them?

In Chapters 4 and 5, we examine the spatial configuration of a character by comparing the performance for upright characters and their inverted versions with psychophysics and fMRI. In Chapter 4, we investigated the spatial configuration processing of orthographic objects using Chinese characters. Part of the material in this chapter has been published in *Cortex*. In Chapter 5, we investigate the relationship between the local and the global configuration in a character, by comparing the inversion effect of a character and its components with psychophysics and fMRI. In this study, we used multivariate methods (such as support vector machine, SVM) with fMRI to classify the different representations of character information in the visual cortex and to reveal the neural computations in the visual cortex for the visual word analysis. Finally, in Chapter 6, we discuss how these behavioral data and neural activities fit into the hierarchical processing for visual word form. We also highlighted future directions and implications.



Chapter 2 : Spatial summation of visual word form processing

In Chapter 2, we investigate the mechanisms for visual word form perception by measuring the contrast threshold for detecting characters and for discriminating real characters from other character-like stimuli. Our approach is based on two experimental paradigms: The dual task paradigm developed by Thomas (1985a, 1985b) and the spatial summation paradigm by Tyler and Chen (2000, 2006).

In the dual task paradigm of Thomas (1985a), the observer had to make two responses in each trial. In his experiment, there were two intervals in each trial. One interval contained the noise mask alone while the other contained the noise mask plus one of the two possible targets. The observer had to first decide which interval contained the target (detection) and then decided which one of the two targets was shown (identification). He found that while the observers' performances in the two tasks were similar, the threshold in the identification task was higher than that in the detection task. His result implies that the observers may rely on the response of one channel to detect the target and the maximum response of several channels to identify the target. That is, the difference in the two tasks reflects the properties of the mechanism that determines the performance in these tasks (Thomas, 1985b).

The spatial summation paradigm has been used to estimate the spatial extent of the receptive field of pattern detectors (Barlow, 1958; Tyler & Chen, 2006; Wilson & Wilkinson, 1998). In a typical spatial summation experiment, the task of an observer is

to detect target stimuli of various sizes. In general, the target detection threshold decreases with the target size to a certain value. Beyond this critical size, there is little, if any, further threshold. The spatial summation has been used to test whether there is a specific visual mechanism for object (e.g. face) processing (Tyler & Chen, 2006). According to the assumption of the linear system (Blakemore & Campbell, 1969) and the signal detection theory (Green & Swets, 1966; Tyler & Chen, 2000), the slope of the spatial summation curve reflects the properties of the underlying mechanisms. Thus, the analysis of the spatial summation for characters could help us to explore the factors that affect visual word form perception.

In the current study, we combined the dual tasks (the detection task and the discrimination task) and a spatial summation paradigm to examine the mechanism of visual word forms. In the detection task, the observers were to indicate which one of the two spatial intervals contains a stimulus. In the discrimination task, the observers had to decide which interval contains a real character while the other contained one of the other stimulus types. The discrimination task was similar to the lexical decision task in the linguistic literature. If the two tasks show no difference, then there should be only one mechanism involved in visual word form processing. There will be no hierarchical processing and no visual word form specific mechanism. Otherwise, how the visual performance differs in these two tasks tell us the properties of the mechanism underlying these two tasks.

In general, both the critical size and the slope of threshold reduction also depend on the spatial extent of the target detector. Suppose that the target is sufficiently small and it lies completely with the receptive field of the smallest detector. As the target size increases, it would simply increase the areal overlap between the detector and the

stimulus and generate a proportional increase in the response of the detector. Under the assumption of linear receptive field, the threshold decreases proportionally with the target size and thus has a slope -1 in log-log coordinates. This reciprocal relationship between the threshold and the target size is also called Ricco's law (Barlow, 1958; Baumgardt, 1959). As the size of the stimulus increases beyond that of the smallest detector, the further increase in target size produces no further response increment in that detector. However, the stimulus may produce a response in other mechanisms also monitored by the visual system. Thus, further summation capability can still exist for a mechanism that pooling responses from these monitored channels. Suppose that the number of the monitored channels, and thus the size of the attention window, is the visual system is independent of the target size. The larger target should produce a response in a greater number of channels within the attention window and thus reduces the uncertainty in the system (Tyler & Chen, 2000). Tyler and Chen (2000) showed that the signature of such uncertainty reduction effect is that the target threshold should decrease with target size with a slope of -1/2. Finally, when the target size is even larger than that of the attention window of the system, the threshold would be constant regardless the further increase in target size. Hence, the corner between the proportions of the summation curve with slope -1 and -1/2 denotes the size of the receptive field of the smallest target detector while the corner between those with slope -1/2 and 0, the size of the attention window.

In this study, we used five types of stimuli (real characters, pseudo-characters, non-characters, *Jiagu* characters, and scrambled characters) to examine the role of spatial configurations of characters and how they are grouped for visual word perception. All characters were constructed from two components with left-right

configuration. The real characters had two components arranged in a left-right configuration, with the left component denoting the meaning of a character and the right component providing a clue to the pronunciation of a character. Following the character construction convention, a pseudo-character was constructed by putting a phonological component on the right and a semantic component on the left side. For a non-character, the spatial arrangement of the two components contained in a real or a pseudo-character obeyed character construction convention, the arrangement of the two components in a non-character was not. Thus, while the local structure of the character components was kept, the spatial configuration of the character was destroyed. *Jiagu* characters were ancient texts found on archaeological sites excavated in the 20th century. They have the same left-right configuration as modern Chinese characters, but contain no familiar components. For modern readers, *Jiagu* characters were just a combination of two unintelligible pictures. That is, the non-characters would keep the components while destroy the spatial configuration between them and the *Jiagu* characters would have no familiar components while keep the spatial configuration intact. We also added the scrambled characters as our control stimuli. If there is no specific mechanism for visual word forms, we should find the same contrast thresholds for all character types. If there is a specific visual word form mechanism and the function of this mechanism is to analyze the orthographic rules, the detection thresholds for the real and pseudo-characters should be lower than that for other types of characters and it is more difficult to discriminate the real from pseudo-characters than from others. If the mechanism responses to the familiar components of a character, the detection thresholds for the real, pseudo-, and non-characters should be lower than that for other types of characters and it is more difficult to discriminate the real from pseudo- and non-characters than from others. If the mechanism responses to the spatial

configuration of a character, the detection thresholds for the real, pseudo- and *Jiagu* characters should be lower than that for other types of characters and it is more difficult to discriminate real from pseudo- and *Jiagu* characters than from others.

In addition to the issue of the specific visual word form processing in visual system, we are also interested in whether there may be multiple mechanisms of visual word form perception in different eccentricities. Researchers found that it is much harder to identify a letter when it is presented in the peripheral vision field than when it is presented in the central vision field, even with the same cortical size of the stimuli (Chung, 2002; Legge, Mansfield, & Chung, 2001). They suggested that different lexical mechanisms might be involved in central and peripheral vision during reading processes (Legge, et al., 2001). However, Pelli, et al. (2006) found that observers' ideal performances for detecting a letter or a word was not affected by eccentricity. Thus, in this study, we presented written words at different eccentricities and varied the size of the visual word form to compare the spatial summation of the visual word form mechanism at central and peripheral vision.

2.1 Methods

Observers. Three observers (KCH, LYL, and ST) participated in the detection task and three observers (KCH, ST, and LYY) were involved in the discrimination task. KCH is an author of this paper and the other two observers were paid observers and naïve as to the purpose of the study. All observers are native Mandarin speakers and have corrected to normal (20/20) visual acuity.

Stimuli. Five types of stimuli (real characters, pseudo-characters, non-characters, *Jiagu* and scrambled characters) were used in the experiments. All the characters were made up of two components arranged in the same left-right configuration. The real characters had two components arranged in a left-right configuration, with the left component denoting the meaning of the character, and the right component providing a clue to the pronunciation of the character. All real characters were randomly selected from the 1500 most frequently used characters of the 5,656 frequently used characters in the Academia Sinica Balanced Corpus ("Academia Sinica balanced corpus. (Version 3) [CDROM] ", 1998). The pseudo-characters and non-characters were selected from the norms prepared by Hue and Tzeng (2000). The 60 pseudo-characters and 60 non-characters contained in the norms were rated for "character-likeness" by 20 college students on a 7-point scale (1: most like a real character; 7: most unlike a real character). The means of the pseudo-characters and non-characters were 2.30 and 5.40, respectively.

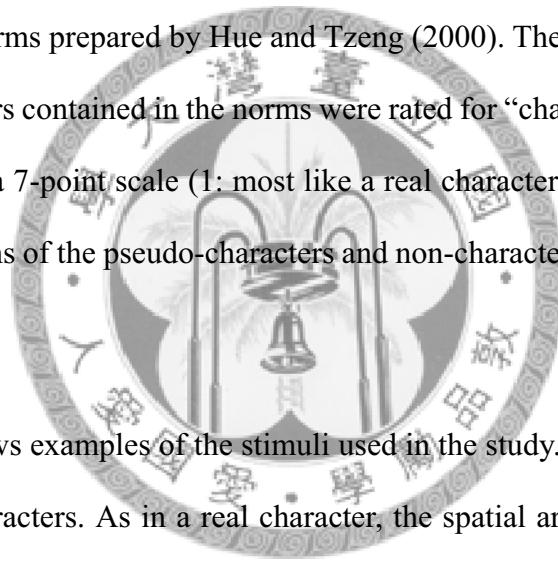
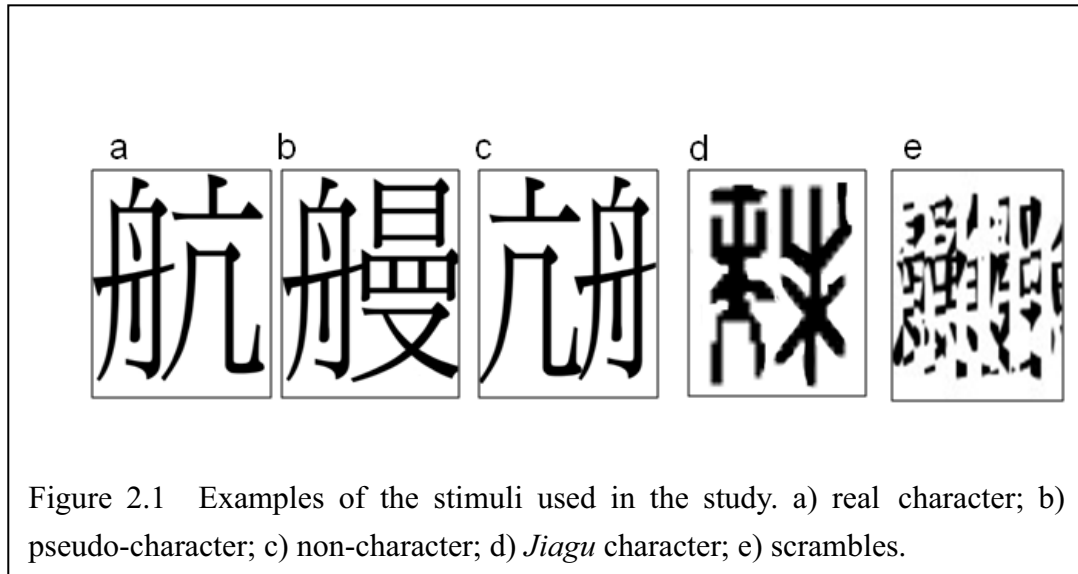


Figure 2.1 shows examples of the stimuli used in the study. Figure 2.1a shows an example of real characters. As in a real character, the spatial arrangement of the two components contained in a pseudo-character obeys the character construction convention and thus makes the character pronounceable to a skilled reader (Figure 2.1b). On the other hand, the positions of the phonetic and the radical of a non-character, in contrast to the character construction convention, were swapped (Figure 2.1c). The *Jiagu* characters were selected from the Digital Archive of the Oracle Bones Rubbing (Institute of History and Philology, Academia Sinica, Figure 2.1d). The scrambled characters were constructed by dividing the image of a character into 16 vertical strips. Then the positions of the 16 vertical strips were scrambled (Figure 2.1e). Each type of

stimuli contained 60 items. All the character stimuli were listed in Appendix 1. The stimuli were presented at 1, 2, 4, and 8 degree away from the central fixation along the horizontal meridian. The visual angles of stimuli were varied from 0.8 to 9.1 degree in half-octave increments.



Apparatus. The stimuli were presented on a HP P1130 (Trinitron 21" CRT) monitor controlled by a RADEON 9800 XT video card on a PC. The monitor resolution was 1280 (H) x 1024 (V) and the monitor input-output intensity function was measured with a light mouse photometer (Tyler & McBride, 1997). The mean luminance of the monitor was 30 cd/m². At a viewing distance of 54 cm, each pixel on the monitor equaled 0.03 degrees in visual angle.

Procedure. We used a spatial 2AFC paradigm to measure the contrast threshold for each type of stimuli at different eccentricities. The Ψ threshold-seeking algorithm (Kontsevich & Tyler, 1999) was used to measure the threshold at a 75% correct level. In the detection task, the experiment was composed of four blocks (real characters, non-characters, *Jiagu* characters, and scrambled characters). In each trial, a stimulus was presented 100 ms and randomly presented on either sides of the central

fixation while the other side was blank. The observers were instructed to indicate which side containing a stimulus. In the discrimination task, the experiment was composed of four blocks (real characters vs. pseudo-characters, real characters vs. non-characters, real characters vs. *Jiagu* characters, and real characters vs. scrambled characters). The target was randomly presented on one side of the fixation, while the other side was one of the other four types of stimuli. The two stimuli of the same size were presented simultaneously for 100 ms. Observers indicated which side contained a real character target.

For both tasks, feedback was provided via an auditory cue for both correct and incorrect trials. The observers were asked to focus on the center fixation point during the entire experiment. There were four threshold measurements and 40 trials for each threshold measurement. The contrast threshold values were the average of four threshold measurements. Each trial had a stimulus presented 1, 2, 4 or 8 degrees away from the central fixation and the presentation order within a block was randomized. The order between blocks was also counterbalanced across observers.

2.2 Results

Figure 2.2 shows the contrast threshold for all types of stimuli at different eccentricities for the detection task. Each column of Figure 2.2 denotes the data from one observer and each row, an eccentricity. In each panel, the blue circle, green square, pink triangle, and red triangle denote real characters, non-characters, *Jiagu* characters, and scrambled characters, respectively. The smooth curves are fits of the model discussed below.

When the target size was small, the detection threshold of a character decreased along with the increase in its size with a slope $-1/2$ (indicated by the dashed curves in Figure 2.2) on log-log coordinates up to a critical size at all eccentricities and for all stimulus types. Beyond the critical size, there was little, if any, improvement in visual word form detectability as stimulus size further increased. Both threshold and the critical size for spatial summation depend on eccentricity and the character type has little effect on detection. While the threshold for *Jiagu* characters may seem slightly lower than the thresholds for the other character types, this could be due to the imbalance of contrast energy for different types of characters. After all, *Jiagu* were originally written with knives on bones and shells. For same overall character size, the strokes of *Jiagu* were wider than those of modern characters that were designed to be written with pens on papers. Hence, at the same Michaelson contrast, the contrast energy for *Jiagu* was still greater than other types of characters. This result suggests that in the detection task, our visual system may analyze local features of characters, and such a mechanism is not affected by familiarity and spatial configurations of a stimulus.

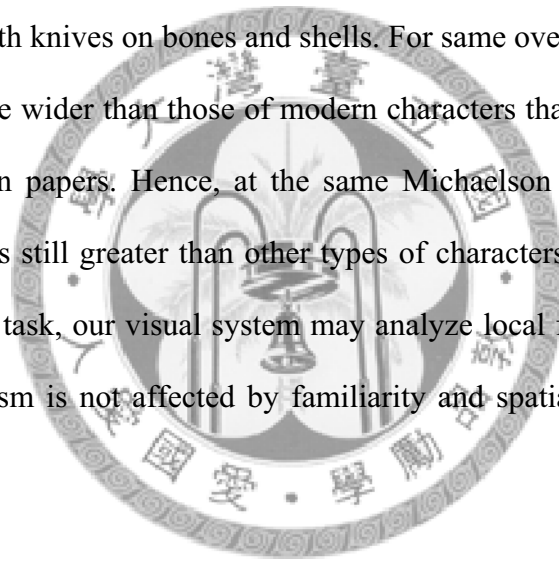
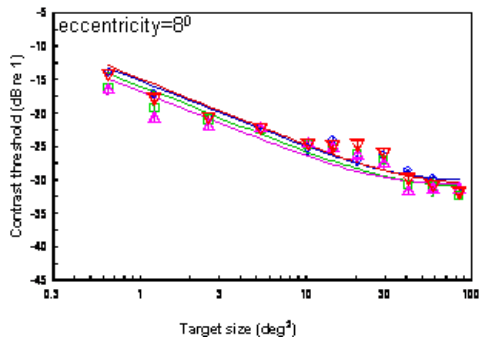
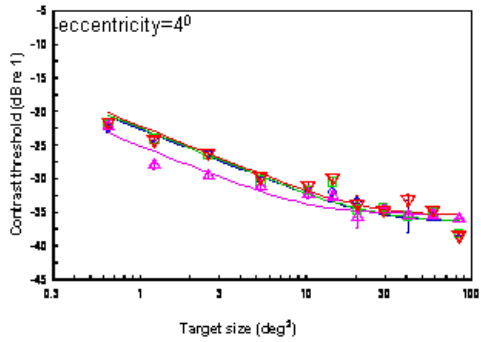
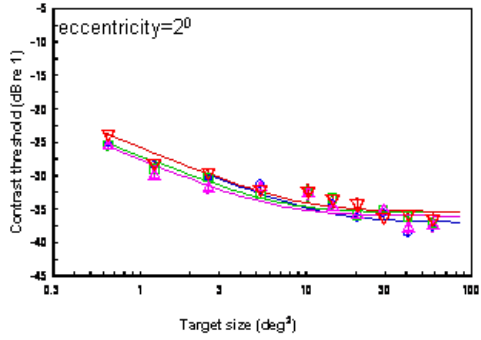
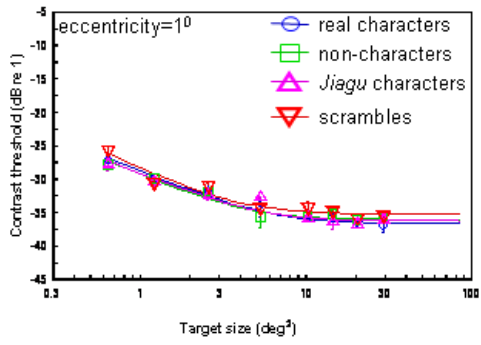


Figure 2.3 shows the contrast threshold for all types of stimuli at different eccentricities in the discrimination task. Each column of Figure 2.3 denotes the data from one observer and each row, an eccentricity. In each panel, the blue circle, green square, pink triangle, and red triangle denote the compared target: pseudo-characters, non-characters, *Jiagu* characters, and scrambled characters, respectively. The smooth curves are fits of the model discussed below. At the same eccentricity, when stimuli were small, the contrast thresholds for the discrimination tasks were higher than those for the detection tasks for all character types. Beyond the critical size, there was no

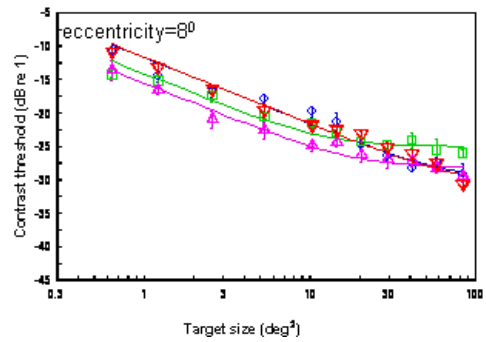
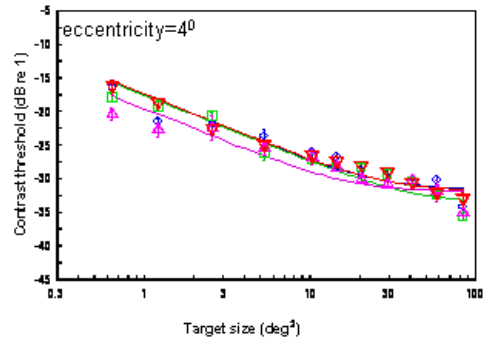
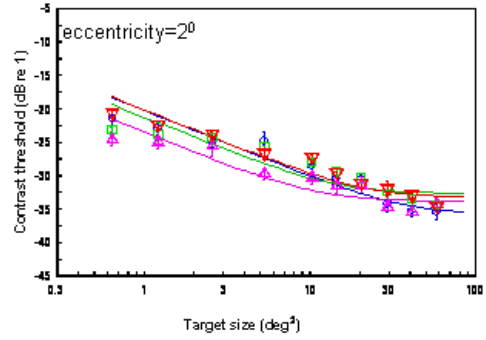
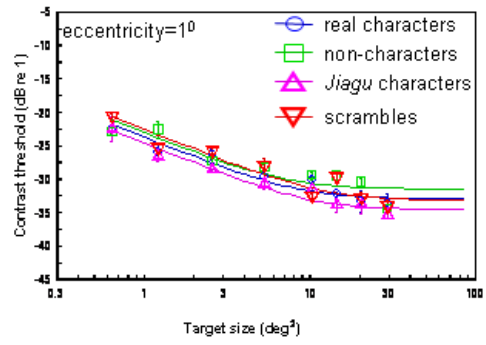
difference in the threshold for the detection and discrimination task. This suggests there is a specific mechanism for visual word form processing.

The discrimination threshold decreased with target size, with a slope -1 (indicated by the dashed curves in Figure 2.3) up to a critical size that was dependent on stimulus type and eccentricity. At the same eccentricity and size, the contrast thresholds for discriminating real and pseudo-characters were the same as those for discriminating real and non-characters. The thresholds were higher, however, than those for discriminating real characters from *Jiagu* characters and scrambled characters. When the target size was small, the threshold for discriminating real and pseudo- or non-characters was increased about 7 dB, or 110%, over that for discriminating real and *Jiagu* characters or scrambled characters. The critical size was largest for discriminating real characters and pseudo- or non-characters and smallest for discriminating real characters and *Jiagu* characters or scrambled characters. Because the difference between pseudo-, non-characters and *Jiagu* characters is that the former contains familiar components while the latter contains no familiar components. Thus, familiarity plays an important role in the discrimination task.

Observer KCH



Observer LYL



Observer ST

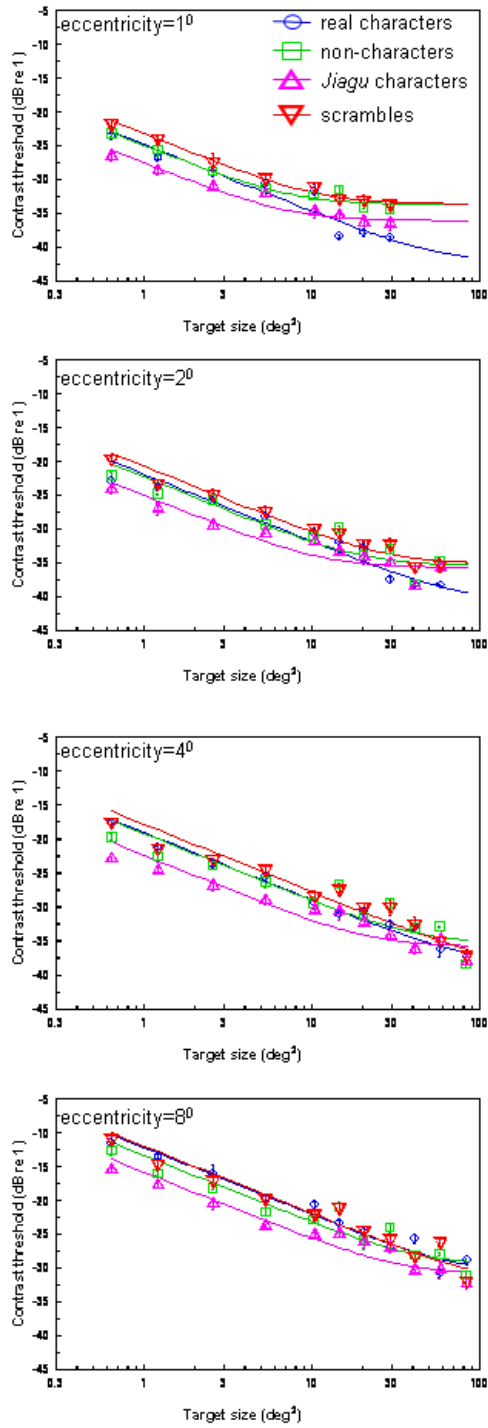
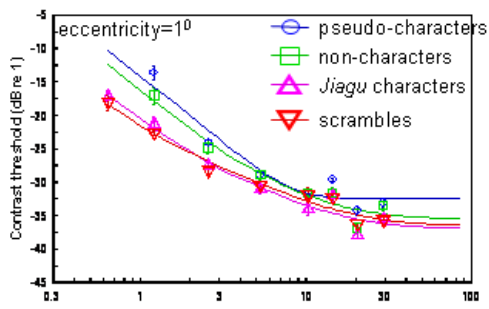
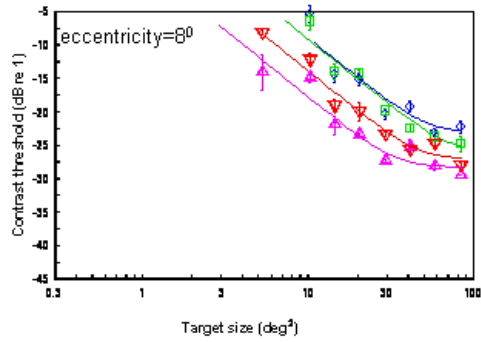
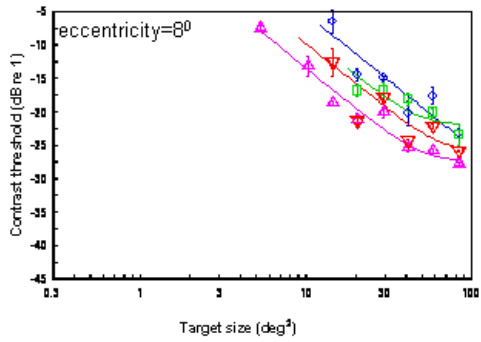
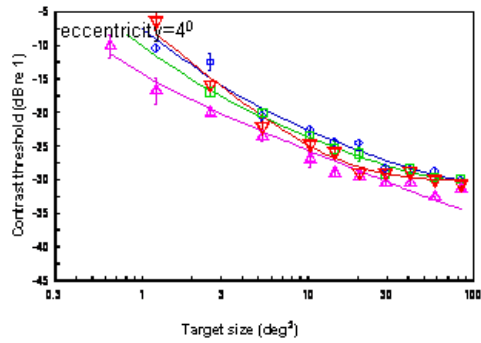
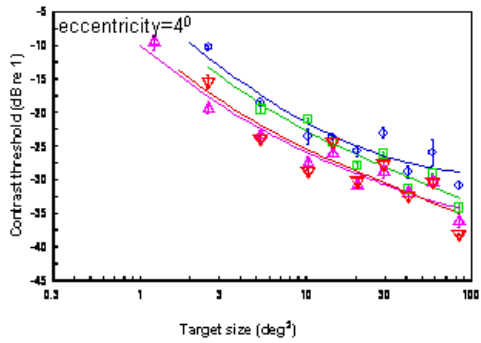
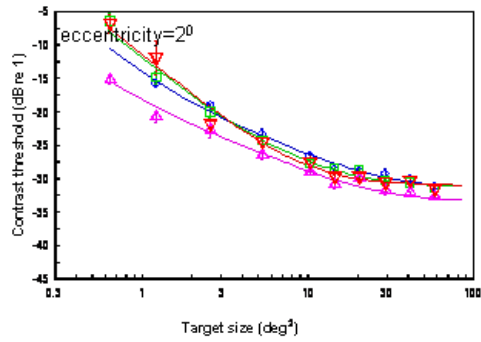
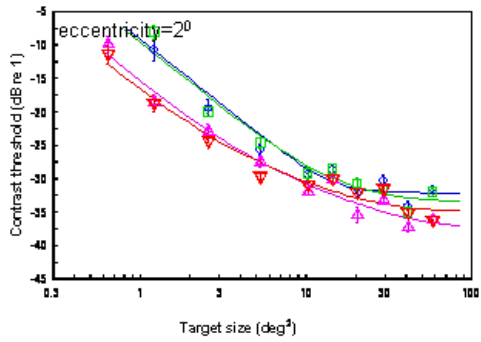
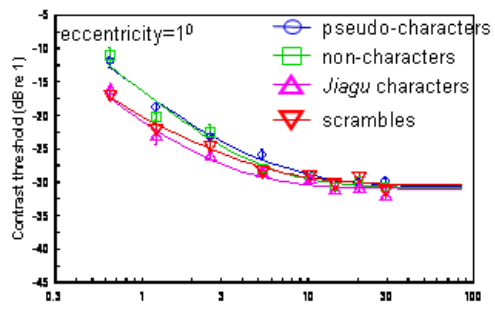


Figure 2.2 The contrast threshold for all types of character-like stimuli at different eccentricities in the detection task. Each panel represents different eccentricities in the detection task. Each row denotes one observer and each column denotes each eccentricity. The spatial summation curves are fits of the model defined in the visual word form. The error bars are one standard error of the means.

Observer KCH



Observer LYL



Observer ST

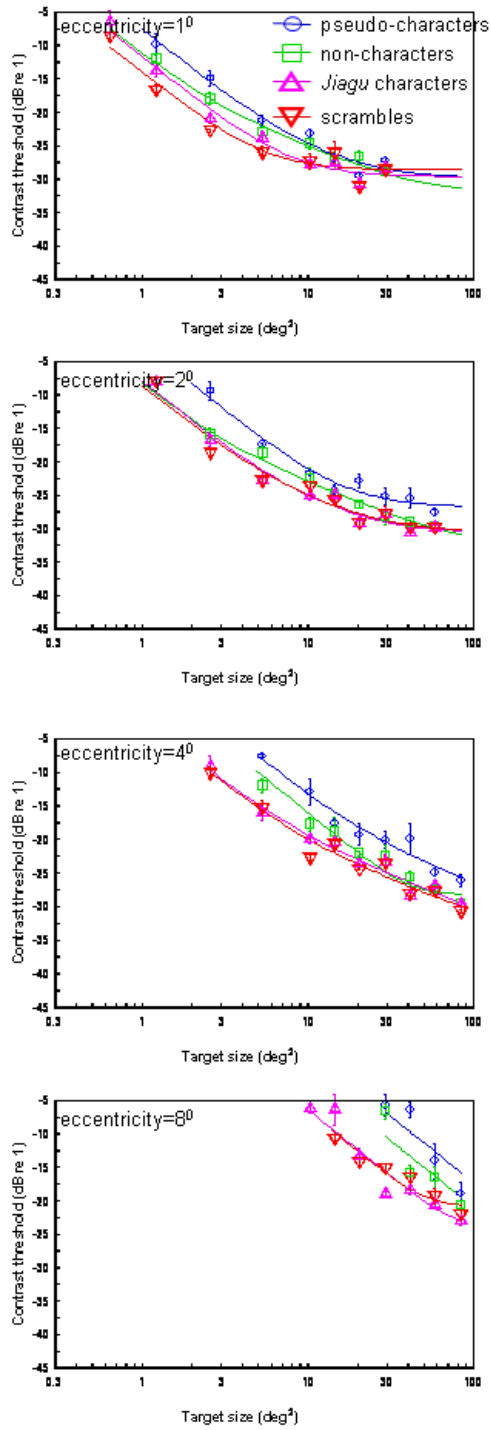


Figure 2.3 The contrast threshold for all types of character-like stimuli at different eccentricities in the discrimination task. Each column denotes one observer and each row denotes each eccentricity. The spatial summation curves are fits of the model defined in the visual word form. The error bars are one standard error of the means.

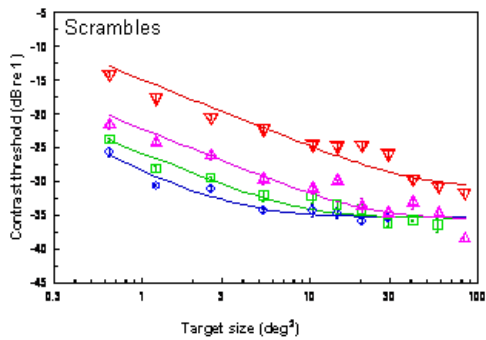
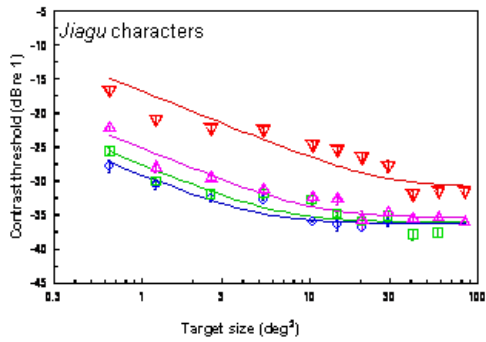
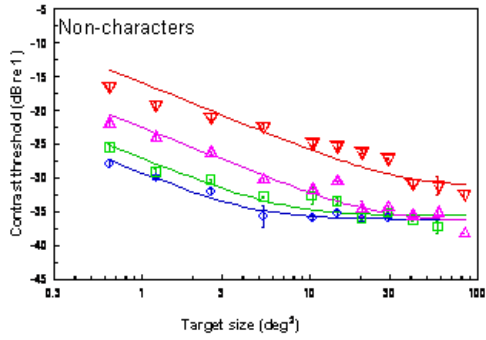
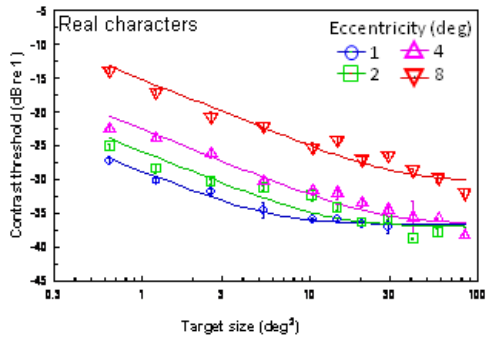
In order to represent the eccentricity effect for detectability, Figure 2.4 replots the data from Figure 2.2. Each column of Figure 2.4 denotes the data from one observer and each column, a stimuli type. In each panel, the blue circle denotes the data collected at 1 eccentricity; green square, 2 eccentricity; pink triangle, 4 eccentricity, and red triangle, 8 degrees. Again, as the target size increased, the threshold decreased to critical size and then leveled out. For all types of characters, the threshold was increased with eccentricity. Furthermore, the performances for all types of characters were significantly reduced in peripheral vision compared with central vision. When the target size was small, the threshold for stimuli presented at peripheral vision (eccentricity of 8 degree) was about 10.3 to 11.2 dB, or 3.3 to 3.6 times greater than that at central vision (eccentricity of 1 degree). When the target size was beyond the critical size, the threshold for stimuli presented at peripheral vision was about 5 dB to 7 dB, or 1.6 to 2.3 times greater than that at central vision. For all types of characters, the critical size was increased with eccentricity. This suggests that the detectability for visual word forms was decreased with eccentricity.

Figure 2.5 replots the data from Figure 2.3 to show the eccentricity effect of discrimination. Each column of Figure 2.5 denotes the data from one observer and each column, a stimuli type. In each panel, the blue circle denotes the data collected at 1 eccentricity; green square, 2 eccentricity; pink triangle, 4 eccentricity, and red triangle, 8 degrees. As the target size increased, the threshold decreased to critical size and then leveled out. For all types of characters, the threshold was increased with eccentricity. The performances for all types of characters were significantly reduced in the peripheral vision, compared with those in the central vision. When the character size was smaller than 3.2 degrees, the threshold in the periphery (eccentricity of 8 degrees)

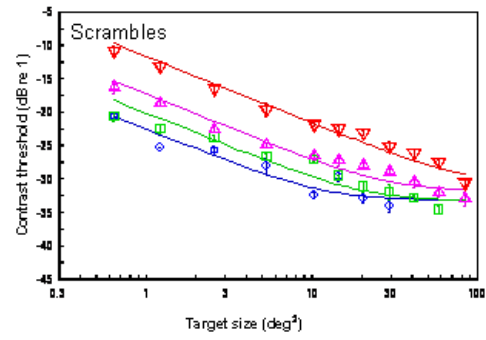
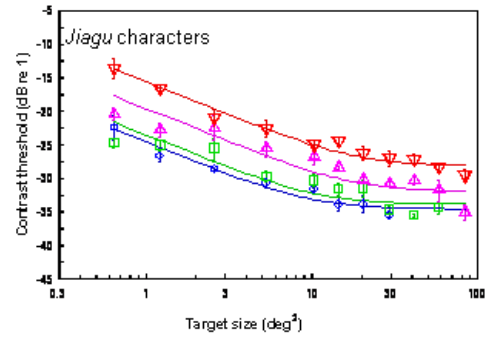
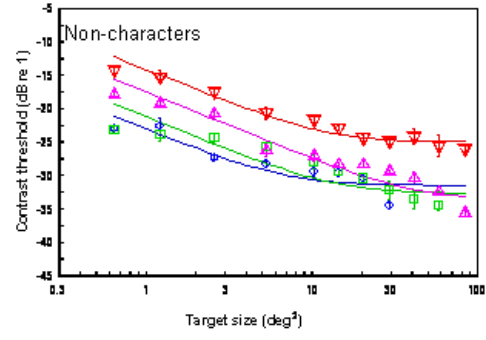
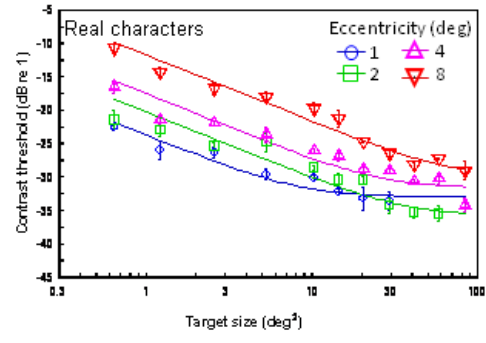
was larger than -5 dB for all types of targets. That is, observers could not distinguish between real characters and other types of characters in the periphery vision when the stimuli were small. The result shows that character size affects the visual word form discrimination in the periphery vision. When the character size was large, the threshold in the periphery was increased by about 8.58 dB more than that in the fovea, that was equivalent to 2.69 times (10 dB, 9.67dB, 6.67dB, and 8 dB for pseudo-characters, non-characters, *Jiagu* characters, and scrambled characters, respectively). This suggests that the discriminability for visual word forms was also decreased with eccentricity.



Observer KCH



Observer LYY



Observer ST

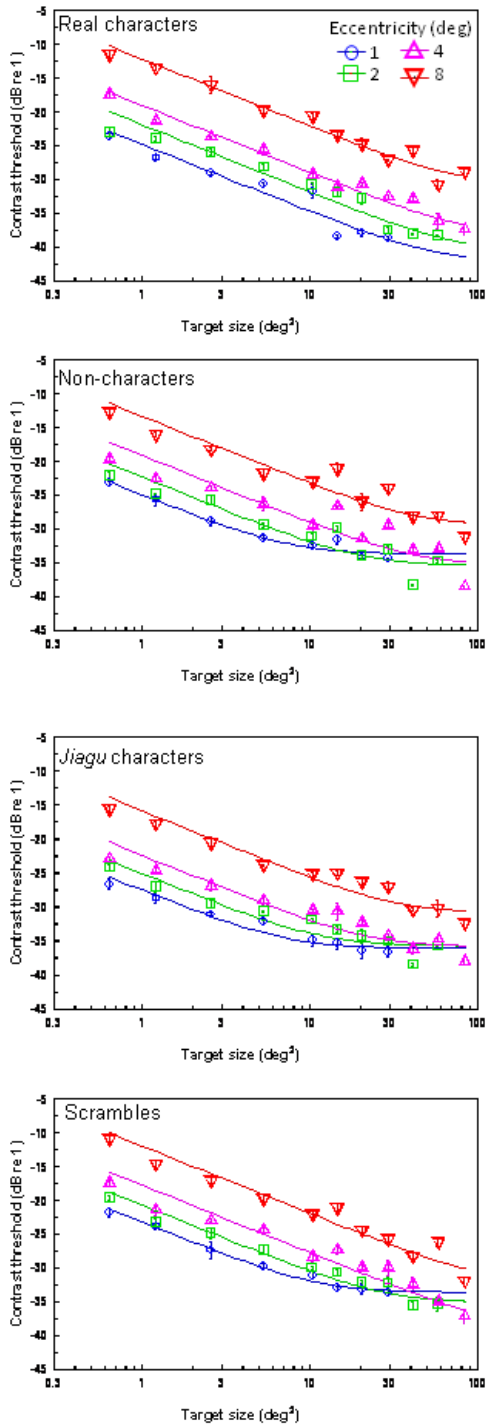
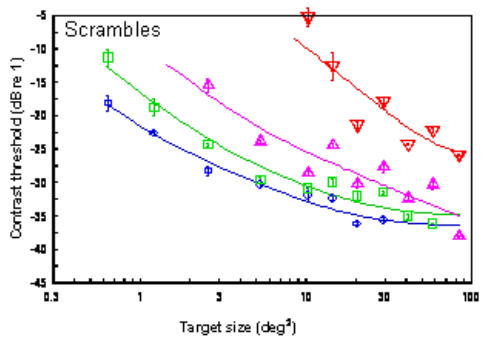
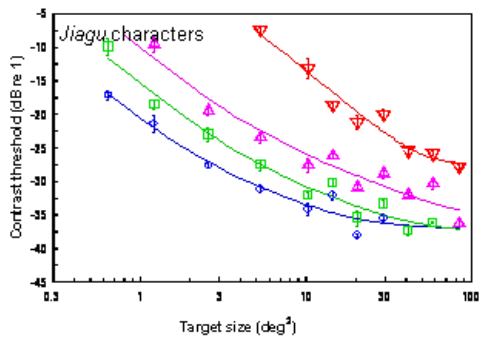
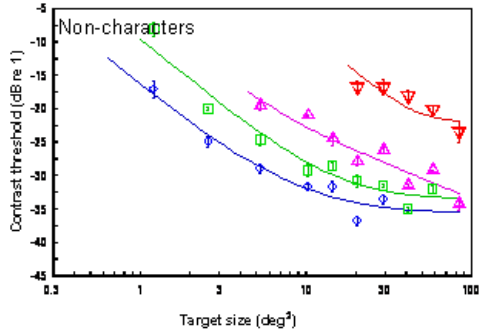
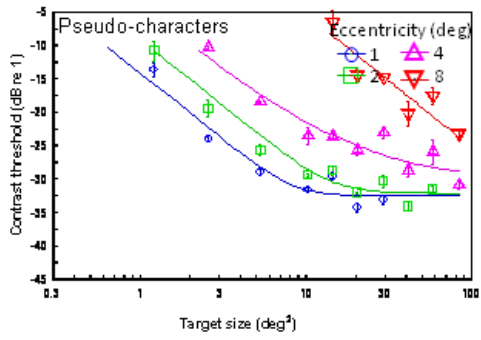
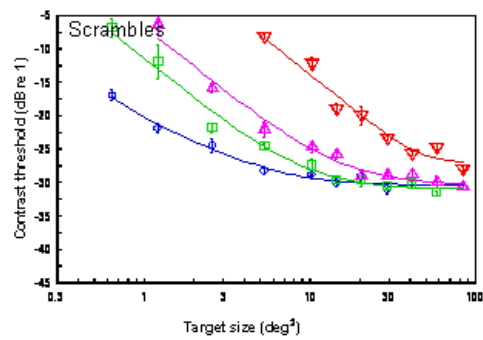
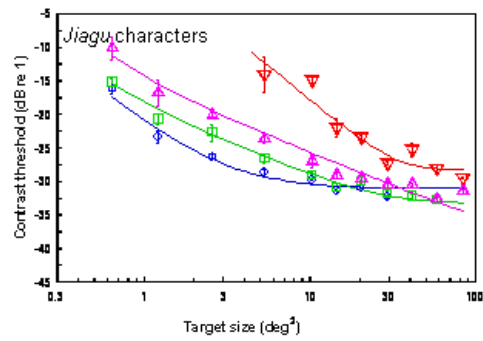
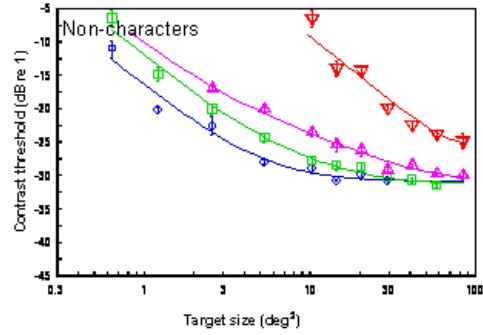
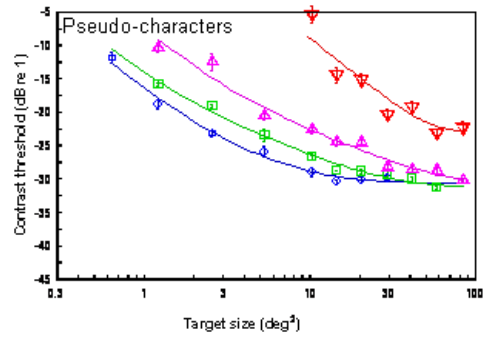


Figure 2.4 The detectability for different eccentricities. Each row denotes one observer and each column denotes each stimuli type. From top to bottom is real characters, non-characters, *Jiagu* characters and scrambles, respectively. The spatial summation curves are fits of the model defined in the visual word form. The error bars are one standard error of the means.

Observer KCH



Observer LYY



Observer ST

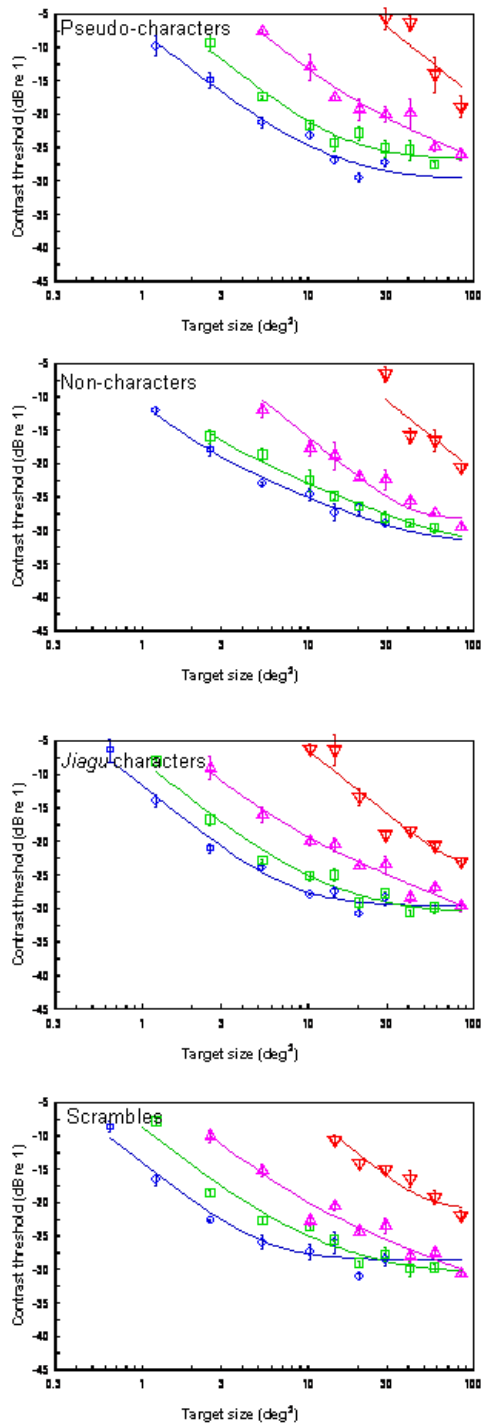


Figure 2.5 The discriminability for different eccentricities. Each row denotes one observer and each column denotes each stimuli type. The spatial summation curves are fits of the model defined in the visual word form. The error bars are one standard error of the means.

2.3 Discussion

Summation model

As we discussed earlier, there are three possible types of spatial summation that can be distinguished by the slope of the summation curves. When the target size is smaller than the receptive field of the smallest visual word form detector, the threshold decreased proportionally with the target size and thus the summation curve would have a slope -1 in log-log coordinates. When the target size is smaller than the receptive field of the smallest visual word form detector, the threshold decreases proportionally with the target size and thus the summation curve would have a slope -1 in log-log coordinates. When the target size is within the attention window of the system, the threshold decreases with the increase in target size, with a slope -1/2 on log-log coordinates. Finally, when the target size is larger than the size of the attention window, the threshold would be a constant regardless the target size. Thus, the summation curve would have a slope 0. We thus combined three possible types of summation with a fourth-power summation operator that provides for smooth transitions between the segments of the curve (Tyler & Chen, 2000) to fit our data. That is, the thresholds θ_i for a character stimulus were fit with an equation:

$$\theta_i = ((s_{1i} \cdot A^{-1})^4 + (s_{2i} \cdot A^{-1/2})^4 + s_{3i}^4)^{1/4} \quad (1)$$

where the subscript i denotes the character type, the s_{1i} , s_{2i} , and s_{3i} are relative weights of the three summation segments. The area of A is computed as ω^2 , where ω is the width of the character stimuli. The smooth curves in Figure 2.2 to 2.5 are model fits.

As shown in Table 2.1, in the detection task, the root mean squared error (RMSE) of the model fit was between 1.29 to 1.53 dB for the stimulus types (1.29 dB for observer KCH, 1.39 dB for observer LYL, and 1.40 dB for observer ST). Although the RMSEs seem larger than the mean standard error of measurement, which ranged from 0.53 to 0.66 dB, this model could account for 98.3 % (98.3% for observer KCH, 98% for observer LYL, 98.6% for observer ST) of the variance in the data. In the discrimination task, the RMSE of the model fit was between 1.12 and 1.37 dB for the stimulus types (1.59 dB, 1.08 dB, and 1.18 dB for observers KCH, LYY, and ST, respectively). This model accounted for 98.8% (99% for observer ST, 98.9% for observer KCH, and 98.6% for observer LYY) of the variance in the data. That is, the model virtually accounts for all the systematic variations in the data.

Detection task	RMSE	SSE
Real characters	1.28	0.61
Non-characters	1.53	0.53
Jaigu characters	1.39	0.66
Scrambles	1.25	0.68
Discrimination task		
Pseudo-characters	1.37	0.81
Non-characters	1.12	0.75
Jaigu characters	1.27	0.70
Scrambles	1.37	0.68

Table 2.1 RMSE and SSE for each type of stimuli.

As we consider the data of detectability for visual word form, the model can be simplified even further. The summation curve at the smaller target size, as shown in Figure 2.2 and 2.4, actually had slope $-1/2$ as indicated by the dashed curves in the figures. Hence, we tested whether a reduced model with $s_1=0$ can fit the data. That is, the thresholds θ_1 was fitted with Equation 2:

$$\theta_1 = ((s_2 \cdot A^{-1/2})^4 + s_3^4)^{1/4} \quad (2)$$

The goodness of fit for Equation 2 is similar to that for Equation 1. The sum of square error pooled across observers increases from 75.33 for Equation 1 to 75.36 for Equation 2 ($F(1,22) < .01, p = 0.999 > .05$).

We consider the data of discriminability for visual word form. The summation curve at the smaller target size, as shown in Figure 2.3 and 2.5, had a slope -1 on log-log coordinates up to a critical size. We test whether $s_2=0$ can fit the data. We refit the threshold θ_1 with Equation 3:

$$\theta_1 = ((s_1 \cdot A^{-1})^4 + s_3^4)^{1/4} \quad (3)$$

The goodness fit for Equation 3 was more worse than that for Equation 1. The sum of square error pooled across observers decreased from 98.58 for Equation 3 to 66.81 for Equation 1 ($F(1,22) = 4.32, p < .05$), suggesting that Equation 1 can better explain the data.

We found that in the detection task, character type had little, if any, effect on detection. This result was actually consistent with previous findings that detectability is based on local feature analysis (Legge et al., 2007; Martelli, et al., 2005; Pelli, et al., 2006; Pelli, et al., 2003; Tan, Hoosain, & Peng, 1995). Based on the theory of threshold

summation behavior (Tyler & Chen, 2000), we estimated the specific mechanism of the visual word form perception. All stimulus types were fitted with a log-log summation slope of $-1/2$ in the detection task. Furthermore, the visual word form detector is limited by the size of the attention window of the system. (Tyler & Chen, 2000).

In the discrimination task, we found that the threshold for discriminating between real and pseudo or non-characters was higher than that for discriminating between real and *Jiagu* characters or scrambled characters. These character specific mechanisms are indifferent to pseudo- and non-characters, suggesting that they ignore orthographic rules. Those mechanisms are indifferent to *Jiagu* characters and scrambled characters, suggesting that they are not for spatial configuration in general. In addition, the slope of the discrimination threshold is close to -1 on log-log coordinates up to a critical size. This result suggested that the discrimination threshold is mediated by a specific mechanism for characters. The discriminability of visual word form information is based on familiar features of an image.

The size of the detection and discrimination mechanism

As discussed earlier in this chapter, the intersection between the proportions of the summation curve with slope -1 and $-1/2$ denotes the size of the receptive field of the smallest target detector while the corner between those with slope $-1/2$ and 0 , the size of the attention window. In the detection task, since there is no -1 slope segment in our summation curve, we can conclude that all our stimuli were larger than the size of the smallest visual word form detector. The width of the attention window, ω_{T1} can be derived from Equation 2 as.

$$\omega_{T2} = [(s_2/s_3)^2]^{1/2} \quad (4)$$

Figure 2.6 shows ω_{T2} averaged across three observers for different character types and eccentricities. As the eccentricity increased, the size of the largest filter also increased, regardless of stimulus types. The filter size in the periphery was twice as large as that in the fovea.

In the discrimination task, both parameter s_1 and s_2 play important information in discriminating real characters and other types of characters. We translated from Equation 1 into the properties of the size ω_{T1} of the smallest filter and the size ω_{T2} of the largest filter by the following relations:

$$\omega_{T1} = [\min(s_1/s_2)^2, s_1/s_3]^{1/2} \quad (5)$$

$$\omega_{T2} = [\max(s_2/s_3)^2, s_1/s_3]^{1/2} \quad (6)$$

Figure 2.7 shows ω_{T1} and ω_{T2} across three observers in the discrimination task. Figure 2.7A shows the smallest filters for all stimulus types. The size of the smallest filter was increased with eccentricity, regardless of stimulus types. Figure 2.7B denotes the largest filters for all stimulus types. The ω_{T2} of *Jiagu* characters at an eccentricity of 4 was un-measurable, thus we excluded this data point. The size of the largest filter was also increased with eccentricity, regardless of stimulus types. For all stimulus types, the parameter ω_{T2} is larger than ω_{T1} in central vision and, beyond the eccentricity of 8 degree, the parameter ω_{T1} is equal to ω_{T2} . This implies that in the periphery, there is a specialized channel for processing visual word form, regardless of stimulus type.

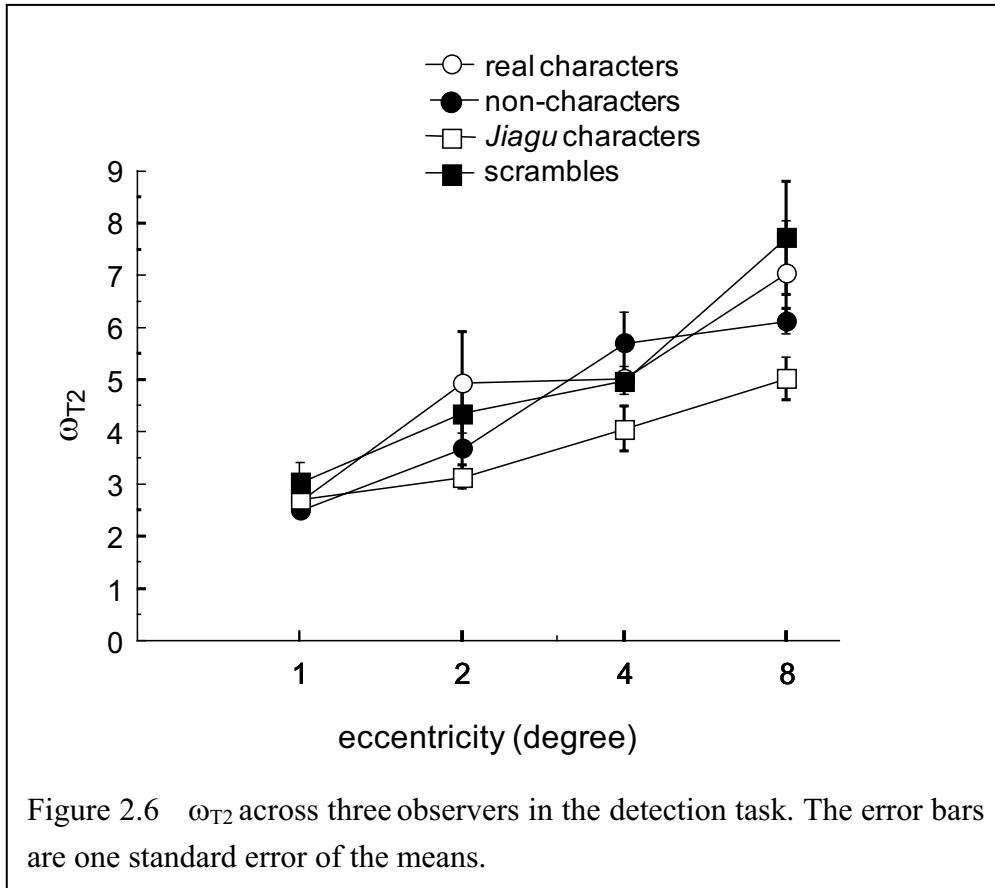
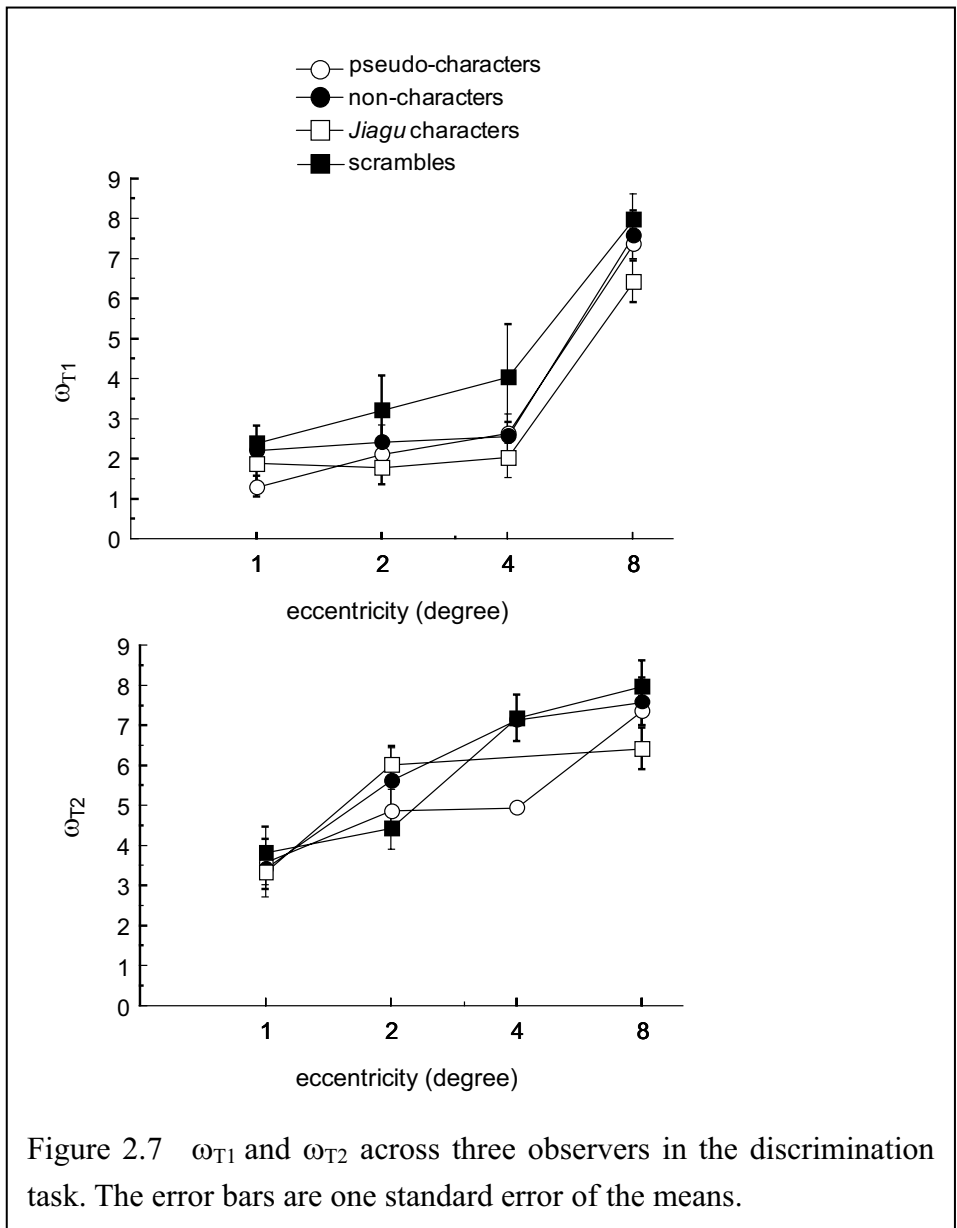


Figure 2.6 ω_{T2} across three observers in the detection task. The error bars are one standard error of the means.





The eccentricity effect

Eccentricity and visual word form size are critical factors in determining reading performance (Battista, Kalloniatis, & Metha, 2005; Chung, Mansfield, & Legge, 1998). To reveal the relationship between eccentricity and visual word form size, we measured

the cortical magnification factors for visual word form processing with Equation 7. The relationship between the critical size ω_T and eccentricity E can be described by

$$\omega_T = \omega_{T0} * (1 + E/E_2), \quad (7)$$

where E is the eccentricity of the stimuli, ω_{T0} is the estimated critical size at the fovea and E_2 controls the cortical magnification factor. ω_T and E_2 were free parameters to be estimated. The E_2 parameter is used to represent the rate of change of the variable of interest as a function of eccentricity (Chung, et al., 1998; Levi, Klein, & Aitsebaomo, 1985; Toet & Levi, 1992). We found the cortical magnification, E_2 , is 0.82 degree visual angle, fitted detectability for visual word forms. From fovea to periphery, the critical size was increased 4.23-fold. The results were consistent with the finding of Levi, Klein, and Aitsebaomo (1985), in which the E_2 was 0.68 for identifying the acuity of alphabetic words. In addition, the critical size increases with eccentricity with E_2 0.22 in the discrimination task. The target size in the periphery was 5.32 times greater than in fovea. The E_2 value in the discrimination task was lower than that in the detection task, implying that the variable changes more quickly with eccentricity for discriminating visual word form than that for detecting visual word form (Chung, et al., 1998).

Crowding effect

Crowding, generally defined as the impairment of target detection with neighboring stimuli, is known as interference between objects, such as faces or written words (Bouma, 1970; Chung, Levi, & Legge, 2001; see review by Levi, Klein, & Chen, 2008). In this study, we demonstrated the crowding effect for detecting and discriminating visual word form. In the detection task, the thresholds for all types of

character stimuli were decreased with increases in the target size. Furthermore, the contrast threshold was greater in the periphery (eccentricity of 8 degrees) than in the fovea. The results seem inconsistent with the findings of Levi, Hariharan, and Klein (2002), and Pelli, Palomares, and Majaj, (2004), who stated that letter detection was not affected by crowding. This inconsistency could be due to the different mechanisms of the “external crowding” and the “internal crowding” (Martelli, et al., 2005; Zhang, Zhang, Xue, Liu, & Yu, 2009). In Pelli et al.’s (2004) study, crowding effect was measured by the critical spacing among letters in the periphery, which is termed as “external crowding”. In our study, the detectability for visual word form in the periphery was affected by the interaction of features within a character, which is termed as “internal crowding”, similar to Zhang et al.’s (2009) finding. A possible explanation for the internal crowding is the interference of mean components in a character (Zhang, et al., 2009). However, we need more studies to understand the mechanism of internal crowding for character perception.

Another important finding is that crowding impaired not only the detectability, but also the discriminability for characters. The crowding effect was greater in the discrimination task than in the detection task. In the periphery, when the target size was smaller than 3.16 degrees, observers could not discriminate between real characters and other types of stimuli, implying that the influence of crowding on discrimination may be based on the combination of features of a character. Furthermore, the crowding for discriminating between real characters and pseudo or non-characters was larger than that for discriminating between real characters and *Jiagu* or scrambled characters. This result suggests that familiar components of a character could interfere with the readability of characters.

2.4 Summary

In this study, we investigated the performance for the detection and the discrimination tasks of displayed visual word form as a function of visual word form-size and retinal eccentricity. We measured the contrast detection threshold for various types of visual word form stimuli (real, non-, *Jiagu*, and scrambled characters) at different sizes and presented at different eccentricities. The detection thresholds for the same stimulus size and the same eccentricity were the same for all types of stimuli. When the visual word form-size is small, the detection threshold of a character decreased with the increase in its size, with a slope of $-1/2$ on log-log coordinates, up to a critical size at all eccentricities and for all stimulus types. This result suggests that detection is mediated by local feature mechanisms, with no specialization for Chinese characters. At the same eccentricity, the contrast thresholds were higher in the discrimination task than that in the detection task, suggesting a different mechanism involving in these two tasks. The contrast thresholds for discriminating between real and pseudo-characters are the same as those for discriminating between real and non-characters in comparable conditions. This result suggests that the visual processing of real, pseudo-, and non-characters may involve same channels. The thresholds were significantly higher than the thresholds for distinguishing real characters from *Jiagu* characters and scrambled characters. The results suggested that the discrimination threshold is mediated by well-learned forms of components in a character. Notice that a real, pseudo-, and non-characters contain similar components. However, it is not clear the role of the components in visual word form processing. This issue will be addressed in the later chapters.

Chapter 3 : Retinotopic mapping of visual word forms in the visual cortex

The purpose of this study was to use the retinotopic mapping method with rotating wedge and expanding ring stimuli to investigate whether the visual cortical areas have specific spatial representations for visual word forms. The retinotopic mapping method has generally been used in defining the borders of the visual cortical areas (DeYoe et al., 1996; Dougherty et al., 2003; Tootell et al., 1998). The organization of the early visual areas is well established and can be mapped the visual field along the polar angle and eccentricity. Each retinotopic area corresponds to specific functional properties of sensory input and forms a hierarchical organization in the visual cortex. Researchers found that the early visual cortical areas respond to certain local features of an object, such as segments, contrast, or contours (DeYoe, et al., 1996; Dougherty, et al., 2003; Tootell, et al., 1998) and the ventral cortical regions respond to complex objects, such as houses, faces, and visual words, and showed the selectivity of different categories (Grill-Spector, et al., 1998; Hasson, et al., 2003; Hasson, et al., 2002; Levy, et al., 2001). Such functional specification in the visual cortex was consistent with the hierarchical processing from early visual areas to higher visual areas.

Although the early visual cortex showed clear retinotopic mapping properties, the spatial topography in the ventral cortical regions is still under debate (Tyler et al., 2005). Researchers argued that the ventral cortical regions were non-retinotopic or less strongly retinotopic (Grill-Spector, et al., 1998; Levy, et al., 2001; Malach, et al., 1995; Tootell, et al., 1998). In contrast, some studies found that the center/periphery

organization of objects within the ventral cortical regions, suggesting a retinotopic property in these regions (Hasson, et al., 2003; Hasson, et al., 2002; Levy, et al., 2001). The inconsistent results may be due to the intrusion of local features in an object that was not balanced in their experiment design. For example, Hasson (2002) only compared the images of faces, buildings and letters at central and periphery. There was much difference in the low level features in those images. It is possible that the activation in the so called higher visual areas may be from a difference in the local features of an image.

In current study we used a 2 (“rotating wedge of characters” vs. “expanding ring of characters”) by 2 (“with a blank background” vs. “with a scrambled background”) design. Combining the results from the “rotating wedge of characters” and “expanding ring of characters”, we could reveal the retinotopic mapping for characters in the visual cortex. Comparing the activation for the rotating wedge with that for the expanding ring of characters between “with a blank background” and “with a scrambled background”, we could identify the cortical regions both for local feature and for whole character analyses. If there is a specific visual region that is sensitive to characters, we should observe that this area shows a greater activation in the “with a scrambled background” condition than in the “with a blank background”. Furthermore, if the traveling wave of characters modulates the neural responses in the visual areas, neurons in these areas should respond preferentially to the characters in a localized part of the visual field. On the other hand, if the neurons in a region respond uniformly to all spatial positions, we would expect there should be no modulation in response to the characters.

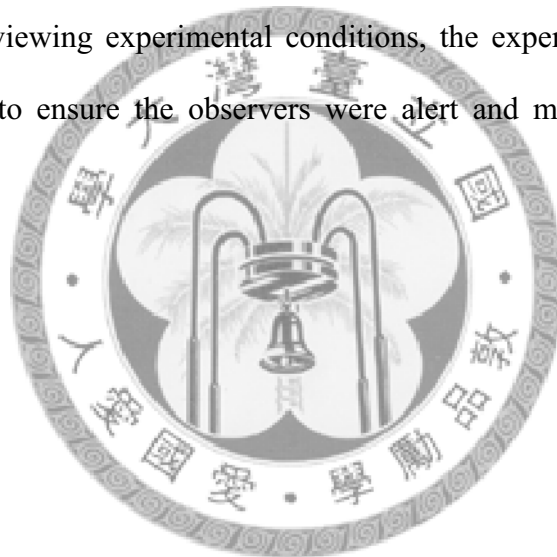
3.1 Methods

Observers Eight healthy, right-handed college students (four males and four females) aged between 22 and 28 years old participated in this study. All the observers were right-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). All the observers were native speakers of Chinese. They were naïve as to the purpose of the experiment and were compensated financially for the hours they devoted to the experiment. Informed consent was obtained from the observers before scanning began. The experiment was approved by the IRB of the National Taiwan University Hospital.

Stimuli. All real characters were the same as those in Study 1. The scrambled characters were constructed first by dividing the image of a character into squares in 4 by 4 grids. The positions of the 16 squares were then scrambled.

Procedure. Four types of conditions were involved in this experiment (see Figure 3.1). There was a 2 (“rotating wedge of characters” vs. “expanding ring of characters”) by 2 (“with a blank background” vs. “with a scrambled background”) design. In the “rotating wedge of characters” condition, characters were presented in one quadrant visual field at a time and smoothly rotated clockwise about the fixation point. In the “expanding ring of characters”, characters were presented in a ring whose radius moved periodically from fovea to peripheral visual field. The images in the conditions were superimposed on a blank background or a background consisting with scrambled versions of characters. Figure 3.1A shows an example of the stimuli in the “rotating wedge of characters only”. Figure 3.1B shows an example of the stimuli in the “rotating wedge of characters with a scrambled background”. Figure 3.1C shows an

example of the stimuli in the “expanding ring of characters only”. Figure 3.1D denotes “expanding ring of characters with a scrambled background”. In all conditions, the character’s size was scaled by eccentricity: from 0.9 degrees in the fovea to 3.6 degree in 7 degrees in the periphery. The expanding ring had a thickness from 1.8° (fovea) to 6.1° (periphery). Stimuli were delivered with MRVision 2000 goggles from Resonance Technology Inc. At 800 (H) by 600 (V) resolution, the theoretical pixel size of the display was 2.25°. In order to limit the differential brain responses to sensory processing signals and avoid confounding the perceptual responses with motor or decision processing, the observers were asked to maintain the central fixation. When each observer was viewing experimental conditions, the experimenter observed the eye-tracker display to ensure the observers were alert and maintaining the central fixation.



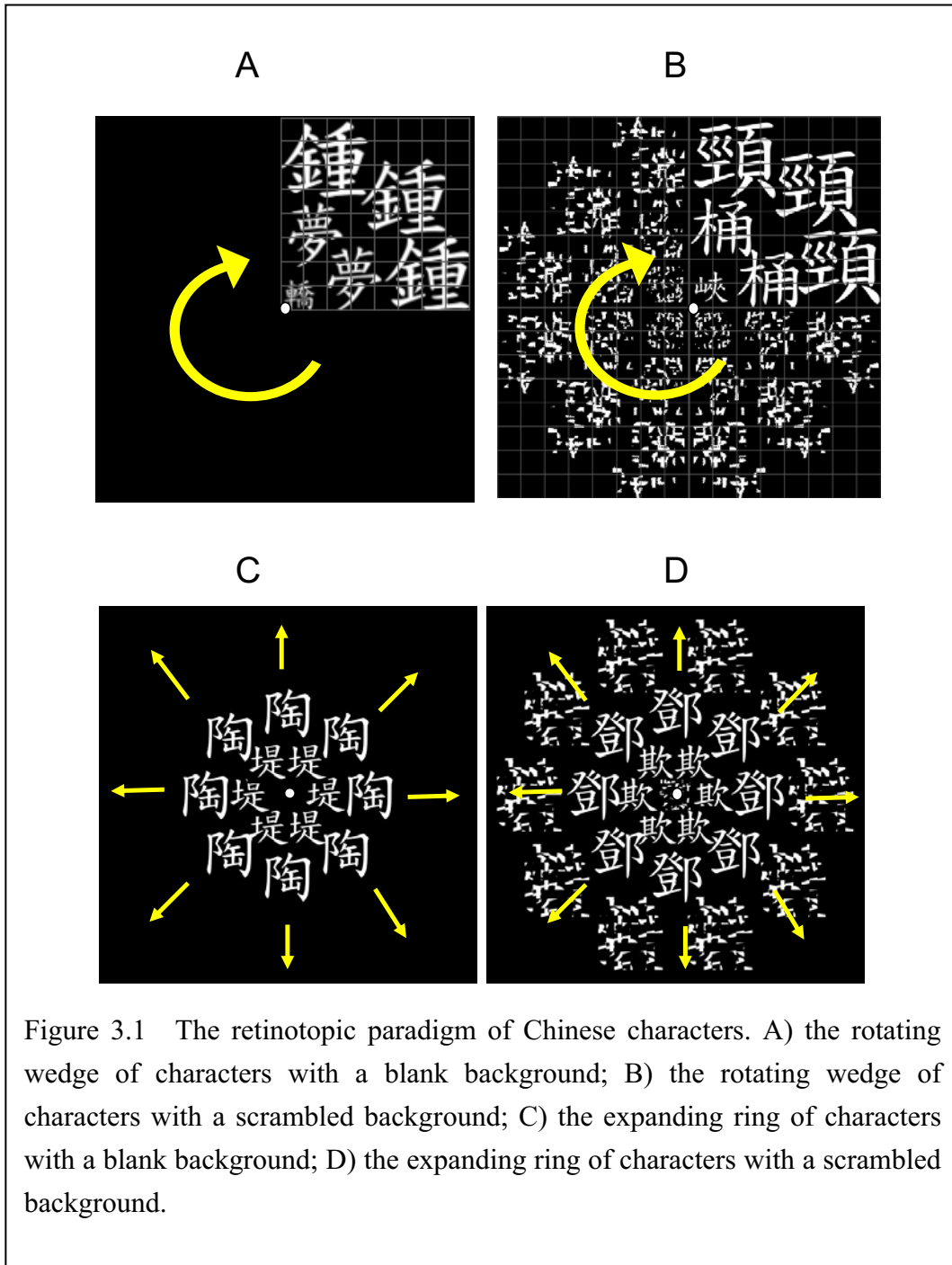


Figure 3.1 The retinotopic paradigm of Chinese characters. A) the rotating wedge of characters with a blank background; B) the rotating wedge of characters with a scrambled background; C) the expanding ring of characters with a blank background; D) the expanding ring of characters with a scrambled background.

Data acquisition. The fMRI images were acquired with a 3T Bruker MRI scanner located at the Interdisciplinary MRI Laboratory at the National Taiwan University. The images were collected in 20 transverse planes parallel to the AC-PC (anterior commissure – posterior commissure) line to cover the occipital cortex. An echo-planar imaging sequence (Stehling, Turner, & Mansfield, 1991) was used to

acquire the functional data (TR = 3000 ms, TE = 33 ms, flip angle = 90°, voxel resolution = 2.34 x 2.34 x 3 mm). Within each scanning session, both functional (T2*-weighted, blood-oxygenation-level-dependent (BOLD)) responses and anatomical images were acquired in identical planes. Each experimental condition lasted 225 s (75 images). The first 9 s (3 images) were excluded from further analyses to avoid the start-up transient. Thus, the data analyzed for each scan spanned 216 s (72 images). The block design was used to measure the cortical response to the stimulus contrasts. Each rotating-wedge or expanding-ring characters were presented for 18 s on each period. There were six periods per run. Each 18 s epoch consisted of 12 presentations of the stimuli for 0.5 Hz followed by a 1000 ms blank period.

Localizers. In addition to the main experiments, all observers participated in localizer experiments in different sessions to identify relevant brain areas reported in the literature. The following four types of localizers were involved in the study. 1) Object localizers. The lateral occipital complex, including the lateral occipital region (LO) and fusiform gyrus (FG), was identified by contrasting BOLD activation with pictures of common objects (objects that can be seen in everyday life) with block-scrambled versions of themselves ($p < .0001$, uncorrected) as developed by Kourtzi and Kanwisher (2001). The borders between LO and FG were separated according to the anatomical structure. 2) Visual word form localizers. VWFA was identified by the differential BOLD activation to real characters versus non-characters ($p < .0001$, uncorrected; Kao et al., 2010). 3) Face localizers. The FFA was identified by the contrast between face images and phase-scrambled face images ($p < .0001$, uncorrected; C. C. Chen, Kao, & Tyler, 2007). 4) Retinotopic localizers. The early visual areas (V1 to V3) were identified with the rotating wedge and expanding ring of

black-and-white checkerboards (Engel, Glover, & Wandell, 1997). The rotating-wedge checkerboards spanned 180° and rotated 30° every 3 s, taking 36 s to go around the display. The expanding-ring checkerboards were presented in a ring whose radius moved periodically from the fovea to the peripheral visual field. The ring had a width of 1.5° whose outer radius expanded from 1.5° to 9° in steps of 0.6° . Each ring was presented for 1 s and the entire sequence of rings was repeated 12 times per scan. The representations of ventral and horizontal visual field meridians were mapped in all observers in order to delineate borders of retinotopic areas.

Data analysis and visualization. A high-resolution anatomical (T1-weighted) 3D MRI volume scan of the entire head was obtained for each observer (voxel size = 1 x 1 x 1 mm). There were three steps to create the flattened cortical representations. First, we used FreeSurfer (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999) to reconstruct each observer's cortical surface based on his or her 3D anatomical brain. Second, the gray and white matters of the cortical surface were segmented with mrGray (Teo, Sapiro, & Wandell, 1997). The information from mrGray could identify the gray and white matter and find the boundary between them. Finally, we used mrFlatMesh (Wandell, Chial, & Backus, 2000) to produce a flattened image of the cortical surface for each observer. The flattened images were fed into the mrVista analysis package (Wandell, et al., 2000) for 3D visualization (see Figure 3.2). The flatmaps were 80 mm radius disks centered at a point near the occipital poles on inflated cortical surfaces. The gray shading denotes the gyral (light) and sulcal (dark) layouts. The differential fMRI activity profile was mapped directly onto the cortical manifold, to allow visualization of the response properties over complete cortical areas. The boundaries of the retinotopic projection areas V1, dorsal V2 (V2d), ventral V2 (V2v), dorsal V3 (V3d),

and ventral V3 (V3v) were established as described in Engel et al. (1997). The areas delineated by red, green and blue borders are the first-tier retinotopic areas V1, V2 and V3, identified by the rotating wedge and the expanding ring of checkerboards, respectively.

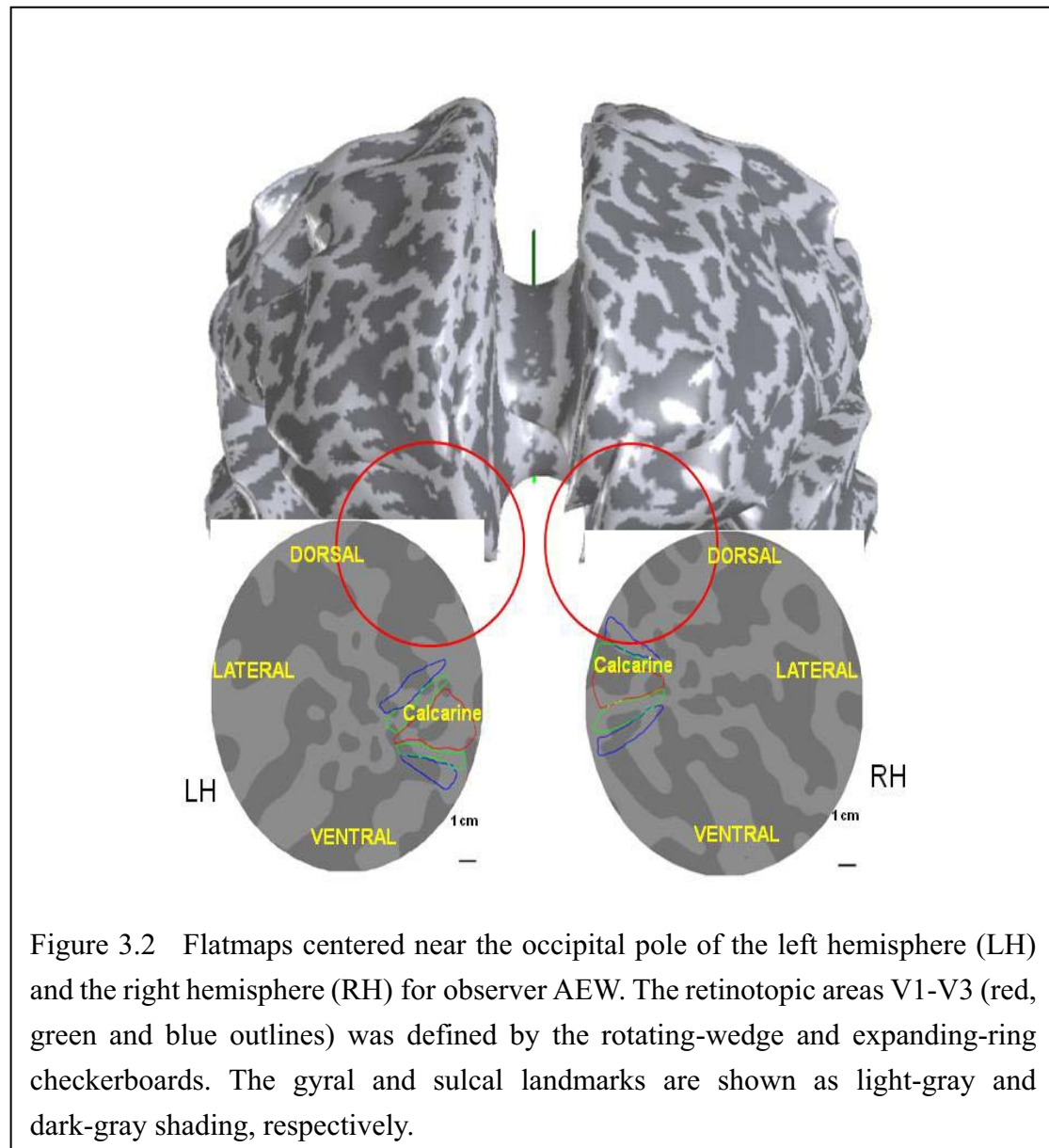


Figure 3.2 Flatmaps centered near the occipital pole of the left hemisphere (LH) and the right hemisphere (RH) for observer AEW. The retinotopic areas V1-V3 (red, green and blue outlines) was defined by the rotating-wedge and expanding-ring checkerboards. The gyral and sulcal landmarks are shown as light-gray and dark-gray shading, respectively.

To correct head-motion artifacts of functional scans, we used SPM2 (Wellcome Department of Imaging Neuroscience, University College London; <http://www.fil.ion.ucl.ac.uk/spm/>) to realign the acquired EPI images. The realigned images, as well as the anatomical images, were then fed into the mrVista analysis package (Wandell, et al., 2000) for coregistration, data analysis, and 3D visualization. No spatial smoothing was used for functional analysis. Statistical analysis of the BOLD activation was based on the spectral correlation between the BOLD activation time series and the experimental sequences (Engel, et al., 1997).

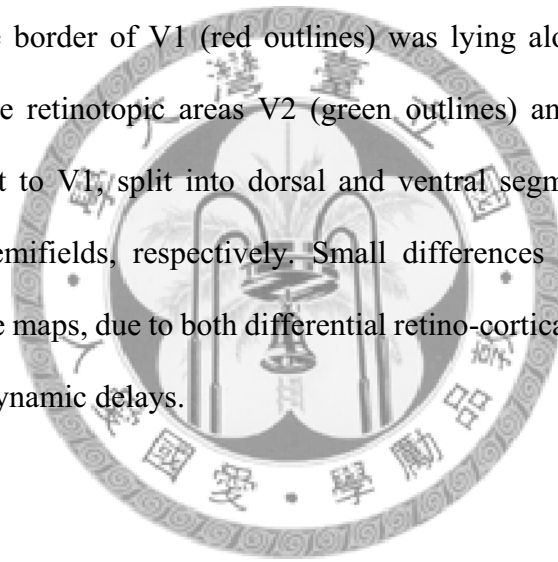
The BOLD response was analyzed by extracting the Fourier fundamental of the time series at every voxel at the stimulus alternation ratio of 1/18 Hz. A statistical correction for multiple occurrences was applied to the criterion for significant response, in terms of the amplitude of the Fourier fundamental. A coherence level of 0.47 provided a significant activation level of $p < .00005$ in each voxel. Responses below this level do not represent significant activation in amplitude terms, although the analysis of the phase vector was implemented under the assumption that that was the best estimate of its phase. The procedure allowed the phase to be used down to a coherence level of 0.3 ($p < .01$ in each voxel) for the ROIs defining visual projection area from the retinotopic scans.

3.2 Results

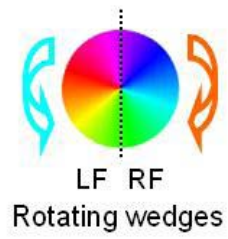
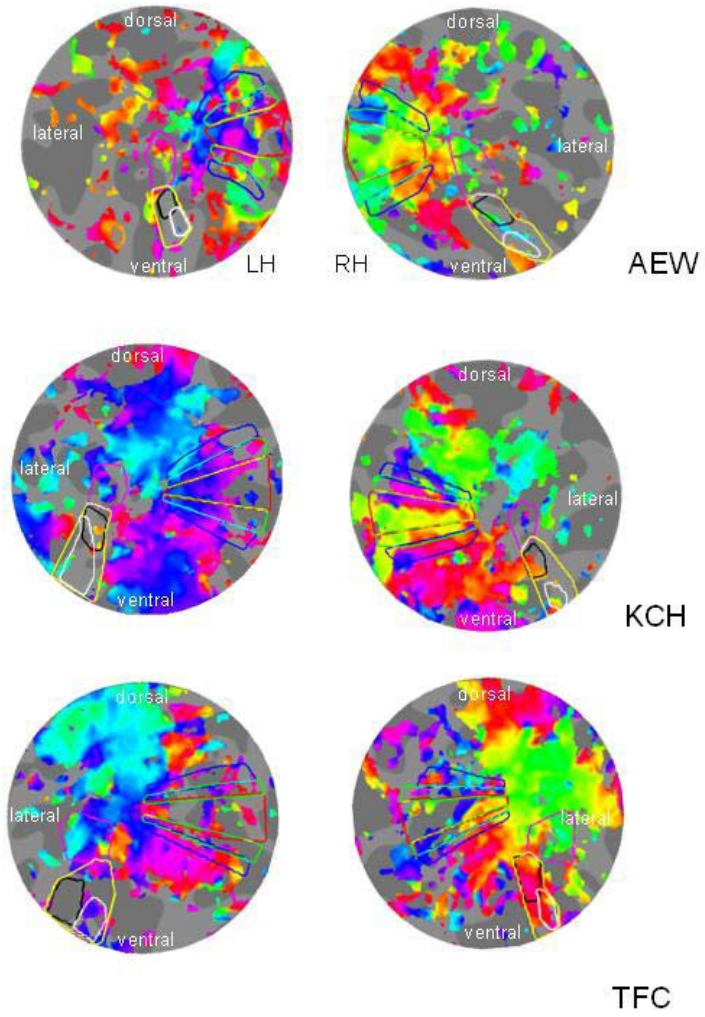
Retinotopic activations for checkerboards

Figure 3.3 showed the retinotopic activations for rotating-wedge and expanding-ring checkerboards for three typical observers (the results of the remaining

observers are presented in Appendix 2). The colors in Figure 3.3 denote the location within the visual field of checkerboards that induced the activity in the occipital cortex. For each retinotopic area, the radial boundaries were defined by the rotating wedge phases, which were coded from cyan in the left and upper visual field to green in the left and lower visual field, and blue in the right and upper visual field to green in the right and lower visual field (Figure 3.3 A). The center to peripheral extent was defined by the expanding ring phases, which was coded from red in the fovea to blue in the periphery (Figure 3.3 B). The retinotopic boundaries obtained in this study were the same as those in previous studies (DeYoe, et al., 1996; Dougherty, et al., 2003; Tootell, et al., 1998). It can be seen that the border of V1 (red outlines) was lying along the fundus of the calcarine sulcus. The retinotopic areas V2 (green outlines) and V3 (blue outlines), which were adjacent to V1, split into dorsal and ventral segments representing the lower and upper hemifields, respectively. Small differences among observers are expected in the phase maps, due to both differential retino-cortical mapping parameters and different hemodynamic delays.



A)



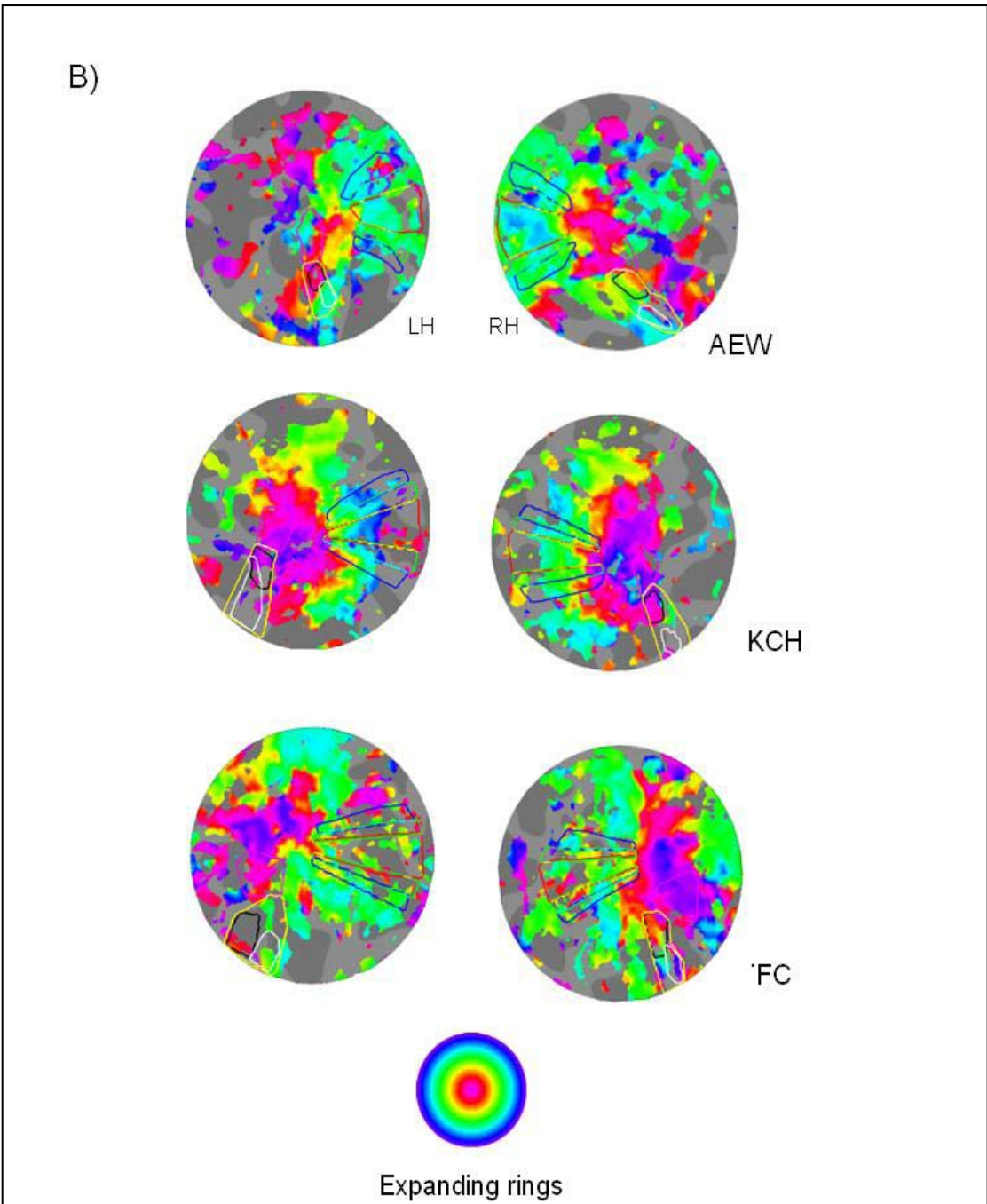


Figure 3.3 Retinotopic activations for the A) rotating-wedge and B) expanding-ring checkerboards for three typical observers. Radial color map segregates the left field(LF) and the right field(RF). Concentric color map separates

Retinotopic activations for characters

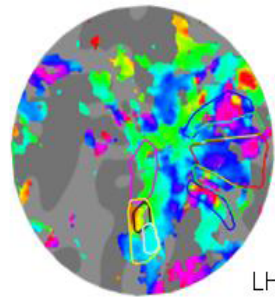
Figure 3.4 showed the retinotopic activations for rotating-wedge (Figure 3.4 A) and expanding-ring characters (Figure 3.4 B) for three typical observers. The colors in Figure 3.4 denote the locations within the visual field of characters that induced the activity in the occipital cortex. The boundaries of early visual areas were defined by retinotopic activations of checkerboards. The magenta, yellow, black and white borders denote the LO, FG, VWFA, and FFA, respectively. The outlines of these areas were shown for reference in the subsequent flatmaps. Traveling waves of characters activated the regions from the occipital pole well into the ventral occipital cortex. The areas responded to the meridional and eccentricity maps and colored according to the color bar. The early retinotopic areas which were defined by checkerboards were also observed in character conditions. Beyond the early visual areas, character stimuli showed a retinotopic property in the LO and FG while checkerboard stimuli did not produce much activation in these areas. This result implies that neurons in the ventral occipital areas were involved in the shapes of visual words rather than low-level features (Grill-Spector, Kourtzi, & Kanwisher, 2001).

We found that the ventral occipital areas responded to characters across the contralateral hemifields and showed the separation between upper and lower visual field representations. That is, this region showed a preference for one portion of visual field compared with another. Both the dorsal and ventral parts of the FG have distinct representations for each quadrant of the visual field. The FG also showed a “center-periphery” property. The FG appears to have a boundary between foveal and

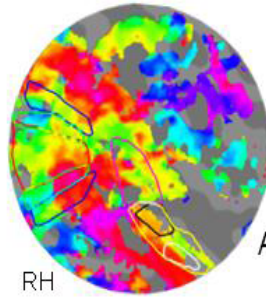
peripheral representations. These results showed that the ventral occipital areas exhibit a degree of retinotopic specificity.



A)

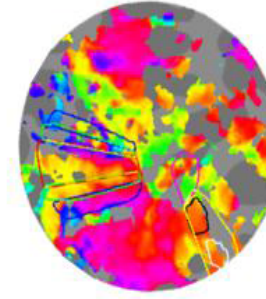
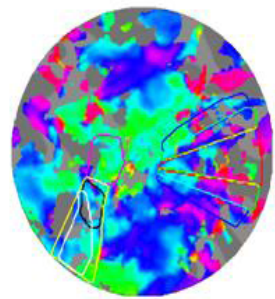
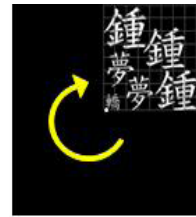


LH

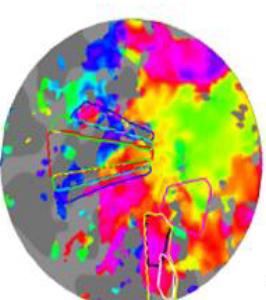
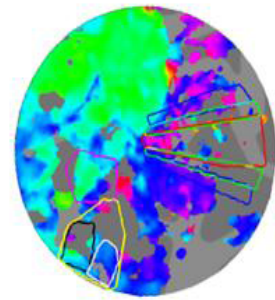


RH

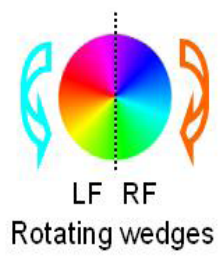
AEW



KCH



TFC



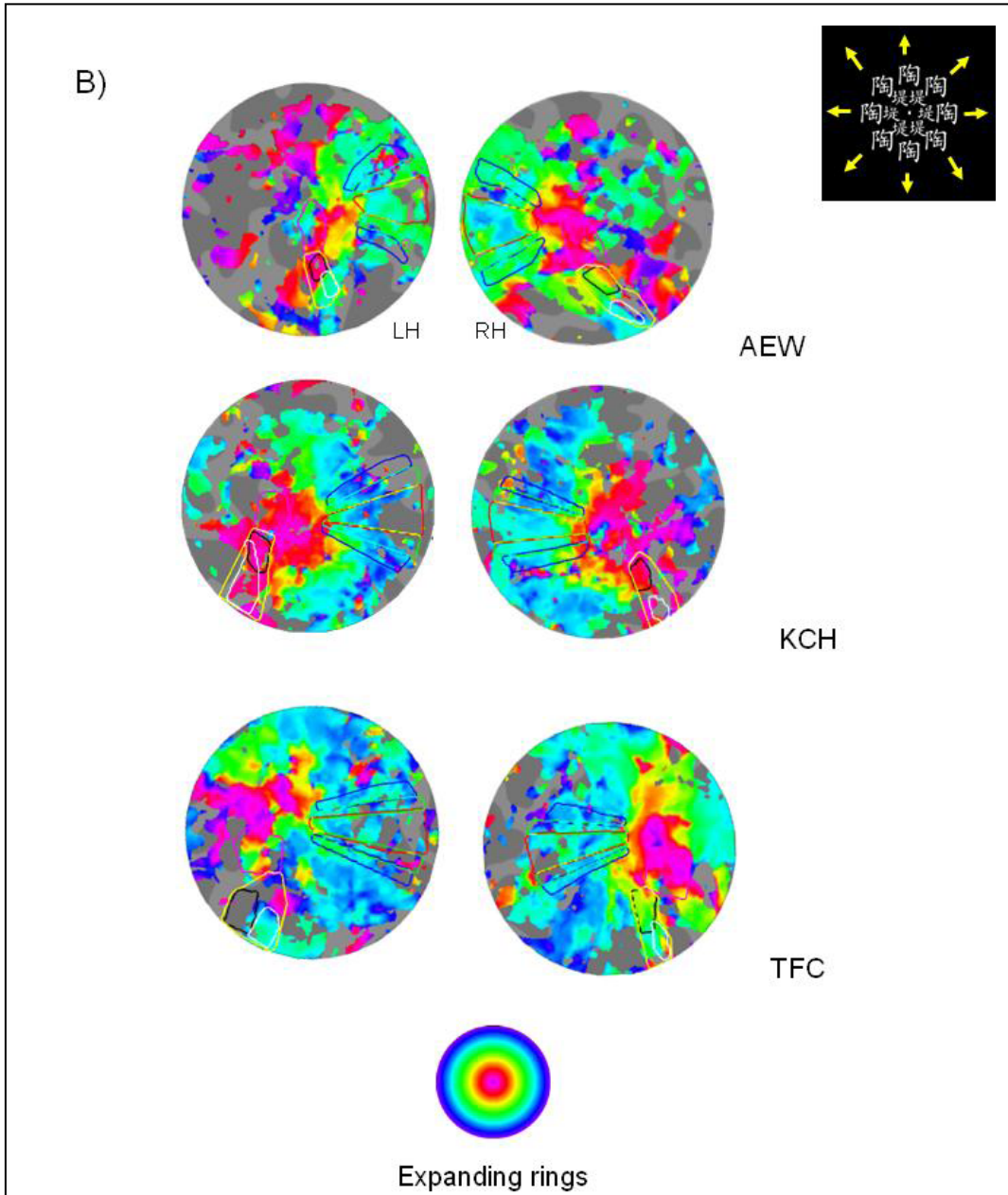
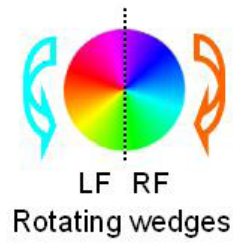
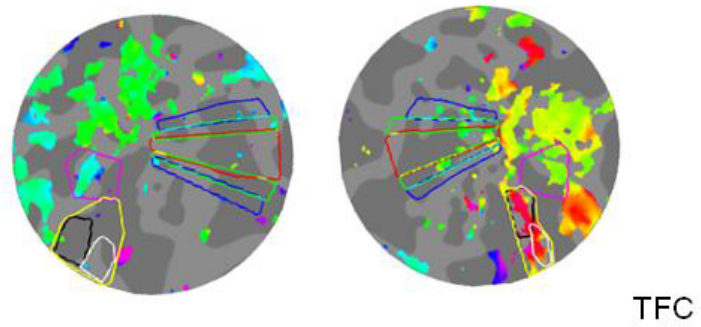
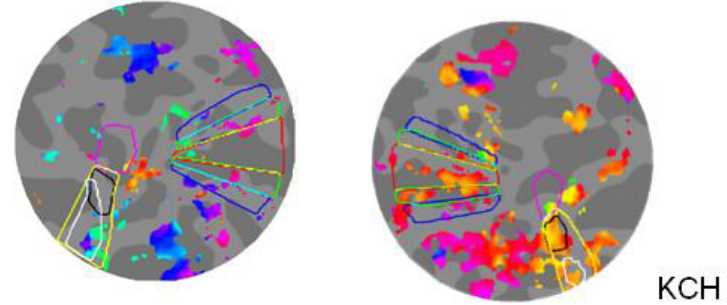
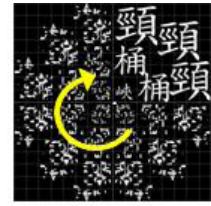
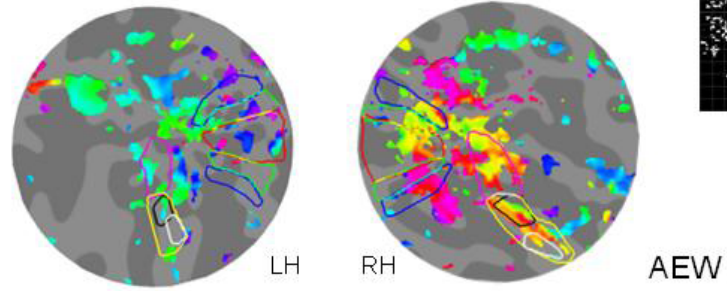


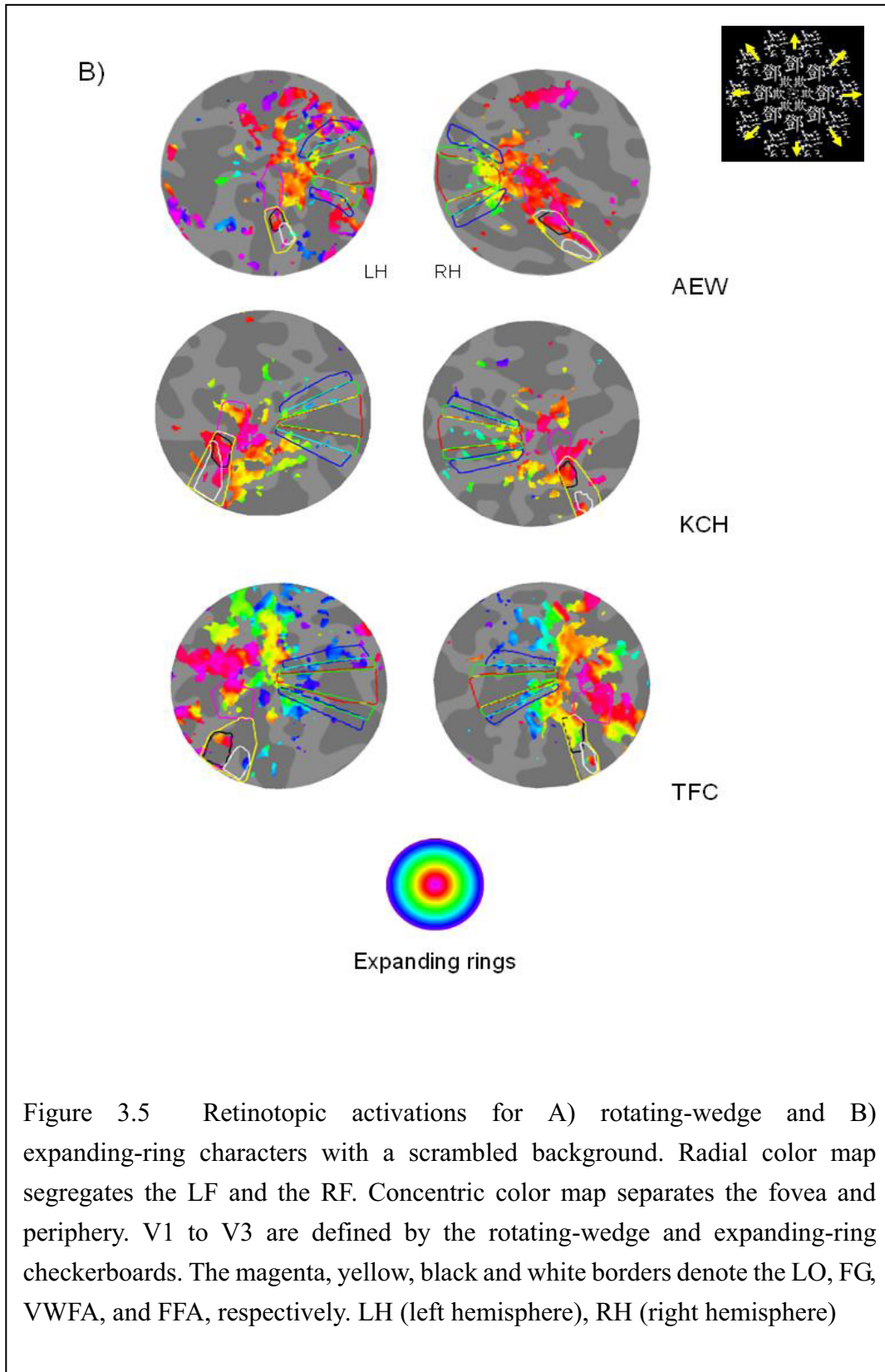
Figure 3.4 Retinotopic activations for A) rotating-wedge and B) expanding-ring characters with a blank background for three typical observers. Radial color map segregates the LF and the RF. Concentric color map separates the fovea and periphery. V1 to V3 are defined by the rotating-wedge and expanding-ring checkerboards. The magenta, yellow, black and white borders denote the LO, FG, VWFA, and FFA, respectively. LH (left hemisphere), RH (right hemisphere)

Retinotopic activations for characters with a scrambled background

Figure 3.5 showed the retinotopic activations for rotating-wedge (Figure 3.5 A) and expanding-ring characters (Figure 3.5 B) with a scrambled background. Contrasted with scrambled characters, the FG responded to characters presented in the fovea, but not in the periphery. More specifically, the foveal selectivity is located within the VWFA, but not in the FFA. The activation of the expanding ring of characters also showed that the bilateral fusiform gyri prefer foveal representation. This result implies that the processing of character recognition may be associated with the central representation in the visual cortex. There is no obvious selectivity, however, for character quadrants in the FG. The activation for characters was only located in the VWFA, not in the FFA. The character-based retinotopic mapping was absent in the FFA. Similarities such as those we describe here have also been noted in other observers. This result was consistent with previous studies that showed the FG has distinct object category representation (Hasson, et al., 2003; Hasson, et al., 2002; Levy, et al., 2001). This suggests a functional segregation in the FG.

A)





3.3 Discussion

In this study, we used the retinotopic mapping method to investigate whether the visual cortical areas have specific representations for visual word forms. Our results showed the activation of characters was extended from early visual areas to the higher visual areas in the occipital cortex. Compared with checkerboards, the character stimuli showed greater activation in the ventral occipital area, especially the FG, suggesting this region is sensitive to the visual words rather than low-level features. The retinotopic representation for character stimuli was in the ventral occipital area, where we found FG is selective for the contralateral visual fields and centre-to-periphery map. This implies that FG is involved in the spatial location information of stimuli. Furthermore, contrasted with scrambled characters, FG showed a preference for characters presented in the fovea. Such foveal selectivity was located in the VWFA, suggesting that VWFA is associated with the processing of central information.

Retinotopic maps in the ventral occipital cortex

Our results clearly demonstrated that the ventral occipital area is a retinotopic organization. We found that FG represented both the contralateral visual fields and upper and lower visual fields. Furthermore, these areas also represented the central to peripheral organizations. The results suggest that the neurons in this region are associated with distinct spatial positions on the retina. It has been known that the ventral part of occipital cortex is activated when participants are viewing objects and is selective for different categories of objects. More recent neuroimaging studies indicate that the ventral occipital area may involve several retinotopic maps. Larsson and

Heeger (2006) reported the two visual field maps (LO1 and LO2), which were between V3d and MT in the ventral occipital area, represented contralateral and upper/lower visual fields. Arcaro, McMains, Singer, and Kastner (2009) reported another two visual field maps (parahippocampal cortex, PHC1 and PHC2) in the ventral occipital area. In addition to these retinotopic areas, we found that the FG represented angular and eccentric maps. These findings support the notion that the ventral occipital cortex is a multiple retinotopic organization (Tyler, et al., 2005; Wandell, et al., 2007).

Each retinotopic area in the ventral occipital area is associated with distinct object tuning properties. The region LO1 was activated more by orientation-selective gratings while LO2 was activated more by objects compared to scrambled objects (Larsson & Heeger, 2006). The regions PHC1 and PHC2 were more responsive to scenes than objects or faces (Arcaro, et al., 2009). The retinotopic map we found in the FG was overlapped with VWFA, implying that this region was selective for visual word processing. Taken together, the ventral occipital area is not a homogeneous area; the populations of neurons may respond to different categories of objects.

Using checkerboards as stimuli, we only found the significant activation in the early visual areas, not in the ventral occipital regions. The higher visual areas seemed not to respond to low-level visual features. Thus, the checkerboards may not be an optimal stimulus to detect the topographic properties of the lateral occipital regions (Tyler, et al., 2005). On the contrary, characters provide suitable stimuli for the investigation of the topography of the ventral occipital surface.

The specific selectivity for characters

Contrasted with scrambled characters, FG showed a preference for characters presented in the fovea. Such foveal selectivity was located in the VWFA, suggesting that the VWFA is associated with the processing of central information. This result supported the notion that object recognition which required central vision is strongly associated with the foveal representations in the brain (Hasson, et al., 2003; Hasson, et al., 2002; Levy, et al., 2001; Malach, Levy, & Hasson, 2002). This result was compatible with the behavioral studies which demonstrated that central vision plays a fundamental role in reading (Legge, et al., 2001). In contrast to the VWFA, the FFA showed no differential activation between characters and scrambled characters. This result suggests that FG has a functional segregation for objects.

One interesting finding is that the right FG is more responsive for the left visual field stimuli when rotating-wedge characters are contrasted with scrambled characters. Behaviorally, the left visual field advantage has been observed in several character recognition tasks, such as lexical decision tasks and naming tasks (Cheng & Yang, 1989; Tzeng, Hung, & Cotton, 1979). They found that the accuracy and reaction times for recognizing characters were better when stimuli were presented in the left visual field than in the right visual field. Cheng and Yang (1989) argued that the left visual field advantage for characters could support the dominance of the right hemisphere in processing characters. The left visual field advantage, which has been reported in behavioral studies, was also found in the activation of the right lateral occipital region in the present study. Activation of the right hemisphere in this region, which was not routinely reported in studies of alphabetic words (Fiez & Petersen, 1998), indicated that

processing Chinese characters which are visually complex involves extensive visual-spatial analyses (Kuo et al., 2001; Tan, Liu, et al., 2001).

Furthermore, our results showed that the left hemisphere responded to the right visual field characters. This result suggests that both right and left lateral occipital regions play important roles in processing characters. Previous studies have shown the left lateral occipital region dominates at visual word perception (L. Cohen, et al., 2000; Dehaene, et al., 2002; Price, 2000) and responds to the lexical information of a visual word. It is likely that the visual word recognition may rely on the cooperation of two hemispheres. We need further studies to test this hypothesis and reveal the nature of the information integration between the two hemispheres.

In addition, the lateral occipital regions showed little or no difference for the upper and lower visual fields, suggesting that the neurons in this area integrate information from these two visual fields. It is possible that receptive field sizes of neurons are sufficient to integrate information across large portions of the visual field.

3.4 Summary

Using retinotopic mapping methodology, we found the activation of characters was extended from early visual areas to the higher visual areas. Character stimuli showed a strong retinotopic property in the FG. This result suggests that the FG is a retinotopic organization. Moreover, character selective activations were identified in the bilateral fusiform gyri. The result of the expanding-ring characters contrasting scrambled characters showed that the bilateral FG prefer foveal presentation. The character-specific area in the FG is different from the FFA. This suggests that the FG is

organized by different functional representations. We found that the perception of characters activated the early and higher visual areas, suggesting that the processing of characters involves the analysis of local features and the global shape. This notion is consistent with the results of object perception studies, in which the visual system analyzes the features of an object and then groups them hierarchically to form object representations.



Chapter 4 : The inversion effect in visual word form processing: the spatial configuration

This chapter has been published as “Kao, C. H., Chen, D. Y., & Chen, C. C. (2010). The inversion effect in visual word form processing. *Cortex*, 46, 217-230.”

Reading may be one of the most well-practiced visual tasks for a person living in a modern society. It is very likely that a reader may have read hundreds of millions of words in his or her lifetime. For instance, skilled Chinese readers can, on average, read about 600 characters per minute (Sun, Morita, & Stark, 1985). Assume that a 20-year-old college student, since the fourth grade, has spent only two hours per day reading at half of his or her top reading speed. This student would have read $300 \text{ char/min} \times 120 \text{ min/day} \times 3650 \text{ days} = 1.3 \times 10^8$ characters in just 10 years. With this amount of practice, our visual system may have developed an efficient way to process word information.

It is suggested that, to perform a well practiced visual task, such as recognizing a face, the visual system tends to analyze the configuration, or spatial relationship among image elements, rather than the image elements themselves (Tanaka & Farah, 1993). At the behavioral level, such configurational processing is best manifested in the inversion effect (Diamond & Carey, 1986; Yin, 1969). In the inversion effect, an observer has more difficulty recognizing an object in a picture when the image is placed upside

down than when the image is in the upright position. Since the image elements in the upright and the inverted images are the same and only the spatial relationship among image elements is changed by the inversion, such impairment of performance implies a mechanism for analyzing spatial configurations whose function is disrupted by the image inversion. Such an inversion effect has been found in recognizing human faces in normal populations (Carey & Diamond, 1977; Leder & Bruce, 2000; Rhodes, et al., 1993; Yin, 1969), recognizing dogs by dog training experts (Diamond & Carey, 1986) and recognizing novel objects by observers trained to identify those objects (Gauthier & Tarr, 1997).

In this study, we investigated the spatial configuration processing of orthographic objects using Hanzi, also called “Chinese characters” in some literature. Hanzi, pronounced as Kanji in Japanese, is a set of characters used in several East Asian written languages. There is behavioral evidence of spatial configuration processing in Hanzi. It showed that skilled readers of Hanzi (e.g., Taiwanese or Japanese college students) tended to sort Hanzi with similarities in global spatial relationships among character components, while non-readers (e.g., American college students) tended to sort Hanzi with similarities in character components (Yeh & Li, 2002; Yeh, et al., 2003).

There may be two ways to analyze spatial configurations in a Hanzi character. First, the visual system may analyze the spatial configurations in a hierarchical way. A character is composed of several character components and each of these components is composed of several strokes. Hence, the analysis of spatial configurations may occur on both levels: one is concerned with the spatial relationship among strokes in a character component, while the other is concerned with the spatial relationship among

components in a character. For the purpose of this discussion, we will refer to the former as the local configuration and the latter as the global configuration. Second, the visual system may directly analyze the spatial relationship of the strokes relative to the whole characters. Hence, character components are just a set of strokes and play only a small role in spatial configuration processing.

To make distinctions among possible types of configuration processing, in this study we use only characters that are semantic-phonetic composites. About 90% of frequently used characters are of this type (DeFrancis, 1984). These characters are composed of two components, arranged in a left-right global configuration (see Figure 4.1a for an example). The spatial configurations of a character were manipulated in the following ways: First, we swapped the position of the left and right components of a character as shown in Figure 4.1b. This manipulation altered the global configuration but left the local configuration intact. For semantic-phonetic composites, this left-right swapping is the same as the construct of “non-characters” referred to in some literature (C.W. Hue & Tzeng, 2000). Thus, to be consistent with previous studies, we will identify the original characters as real characters and the left-right swapped characters as “non-characters”. We then inverted both real characters and non-characters to disrupt both local and global configurations (Figure 4.1c & d).

There are studies using mirror reversed words (Dong, et al., 2000; Poldrack, Desmond, Glover, & Gabrieli, 1998; Proverbio et al., 2007; Ryan & Schnyer, 2007). At first glance, such manipulation is similar to ours. However, those studies focused on skill learning and did not address the issue of the visuospatial analysis of the stimuli. As a result, while they showed extensive occipitotemporal activation in the mirror reading

tasks, these studies failed to analyze or separate the functions of different loci of activation in the occipitotemporal cortex during word processing.

While the inversion effect has been demonstrated in several categories of objects, to the best of our knowledge, it has not been reported for Chinese words or characters. Therefore, we first demonstrated an effect with a psychophysical experiment by showing an impairment of the discrimination performance for inverted real characters, compared with that for upright real characters. In addition, it is reported that the inversion effect may have different properties for foveal and peripheral presented stimuli (McKone, 2004; McKone, Brewer, MacPherson, Rhodes, & Hayward, 2007). We, therefore, also tested the inversion effect at the parafoveal and the peripheral eccentricities. We then investigated the cortical activation for spatial configurations in Hanzi characters. The cortical areas that are sensitive to all spatial configurations should show an inversion effect for real characters. Furthermore, the areas that are sensitive to the global configurations should show the differential activation between real and non-characters. The areas that are sensitive to the local configurations should show an inversion effect for non-characters.

4.1 Methods

Psychophysics

Participants. Twelve right-handed observers (6 males, 6 female) between 20 to 24 years old were participated in this study. All participants were undergraduate or graduate students at the National Taiwan University. They are all fluent in reading Chinese. They were naïve as to the purpose of the experiment and were compensated

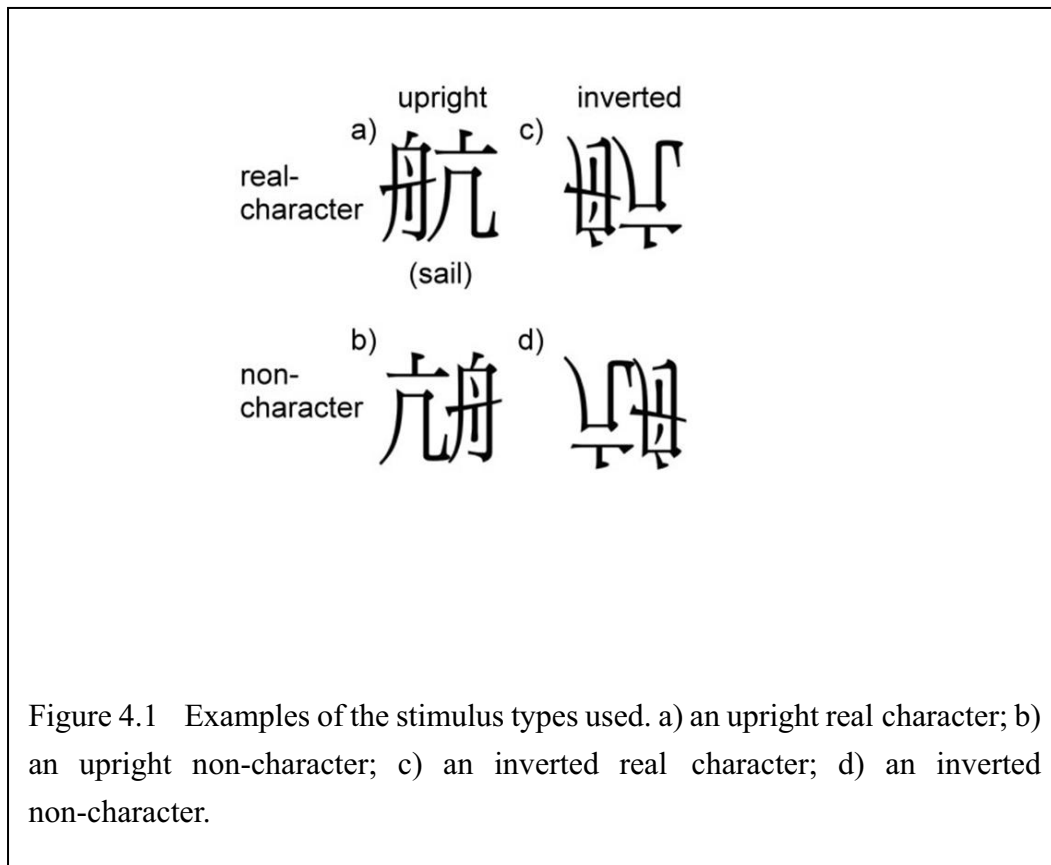
financially for the hours in the experiment. All of them provided written consent for participation in the experiment. Data from one participant was discarded due to a performance near the chance level in all conditions. As a result, there were eleven participants in the final data set.

Apparatus. The stimuli were presented on a HP P1130 (Trinitron 21" CRT) monitor controlled by a RADEON 9800 XT video card on a PC. The monitor resolution was 1280 (H) x 1024 (V). The monitor input-output intensity function was measured with a light mouse photometer (Tyler & McBride, 1997). This information allowed us to compute linear lookup table settings to linearize the output. The mean luminance of the monitor was 30 cd/m². At a viewing distance of 54 cm, each pixel on the monitor had a size of .03° in visual angle.

Stimuli. Figure 4.1 shows examples of stimuli used in this study. All the characters were semantic-phonetic composites with two components arranged in a left-right configuration, with a phonological component on the right and a semantic component on the left side (Figure 4.1a). It is estimated that 90% of the characters used by a typical college student are such compounds (C.W. Hue, 2003). The 40 real characters we used were selected from the 1500 most frequently used characters in the Academia Sinica Balanced Corpus (1998). A non-character (Figure 4.1b) was constructed by interchanging the positions of the phonetic and semantic components of a real character. Thus, a non-character kept the components, and in turn local configurations, of a character, whereas the spatial relationship between components was altered. The 40 non-characters we used were selected from the norms prepared by Hue and Tzeng (2000). All character stimuli were listed in Appendix 3. The inverted

real and non-characters (Figure 4.1c and d respectively) were simply the upside-down versions of their upright counterparts.

The stimuli were presented at either 1° (parafoveal condition) or 4° (peripheral condition) away from the central fixation along the horizontal meridian. The stimuli size was 1.6 deg in the parafoveal condition and 3.2 deg in peripheral condition. The contrast level of the stimuli ranged from -26 to -2 dB in 4 dB step. The unit dB is 20 times the log base 10 of the linear contrast.



Procedure. We used a 2AFC paradigm to measure the proportion of correct responses in a matching task. The experiment was composed of four blocks (upright, inverted real characters, upright and inverted non-characters). In each block, each trial had a stimulus presented in either central or peripheral vision and the contrast threshold of a stimulus was varied with 7 levels. In each trial, two characters of the same type and

orientation were presented. One was on the left and the other was on the right of the fixation point. The two stimuli were presented simultaneously for a 100 ms duration. The participants were asked to press one of two buttons to indicate whether these two characters were the same or not. Feedback was provided via an auditory cue for both correct and incorrect trials. During the experiment, the participants were asked to fixate on the center fixation point. The presentation order within a block was randomized and the order between blocks was counterbalanced across participants. Each block consisted of 560 trials. Thus, each participant was presented with 2240 trials in the psychophysical experiment (4 types of stimuli x 2 eccentricities x 7 contrast thresholds x 40 trials).

fMRI

Participants. Thirteen healthy volunteers (6 male, 7 female) between 22 and 36 years old participated in this study. All the participants were right-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). All the participants were native speakers of Chinese and none of them had participated in the psychophysical experiment. They were naïve as to the purpose of the experiment and were compensated financially for the hours of the experiment. Informed consents of the participants were obtained before scanning. The experiment was approved by the IRB of the National Taiwan University Hospital.

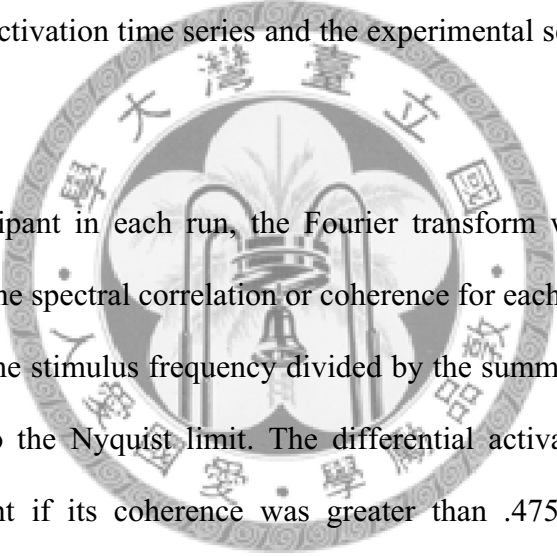
Stimuli. The stimuli were the same as those used in the psychophysical experiment except for the addition of the scrambled versions of the test characters as control stimuli. The stimuli, which had approximately 80% Michelson contrast (-2dB), were presented at the fovea. They were constructed first by dividing the image of a

character into squares in 4 by 4 grids. Then, the positions of the 16 squares were scrambled.

Procedure. We used a block design to measure the differential cortical activations between different types of stimuli. Each of the four experimental runs had an 18s-epoch of the presentation of one stimulus type alternating with an 18s-epoch of the other types. The stimuli were presented randomly within each 18-s epoch. There were six cycles of 36s periods in each run. Each epoch consisted of 18 presentations of the stimulus for 500 ms followed by a 500 ms blank period. Two of the four runs had the real characters alternating with their scrambled or inverted versions respectively, while the other two runs had the non-characters alternating with their scrambled or inverted versions respectively. To keep the participants' attention, the participants were instructed to perform a 1-back matching task in which they had press a key when the stimulus present in the current trial matched that which was presented in the previous trial. In each run, the proportion of matches between stimuli in trial N and trial N-1 was .25.

Data acquisition and analysis. All images were acquired with a 3T Bruker MRI scanner located at the Interdisciplinary MRI Laboratory at the National Taiwan University. A high-resolution anatomical (T1-weighted) MRI volume scan of the entire head was conducted once on each participant (voxel size = 1 x 1 x 1 mm). Within each scanning session, both functional (T2*-weighted, BOLD) responses and anatomical images were acquired in identical planes. The images were collected in 20 transverse planes parallel to the AC-PC (anterior commissure – posterior commissure) line. An echo-planar imaging sequence (Stehling, et al., 1991) was used to acquire the functional data (TR = 3000 ms, TE = 33 ms, flip angle = 90°, voxel resolution = 2.34 x

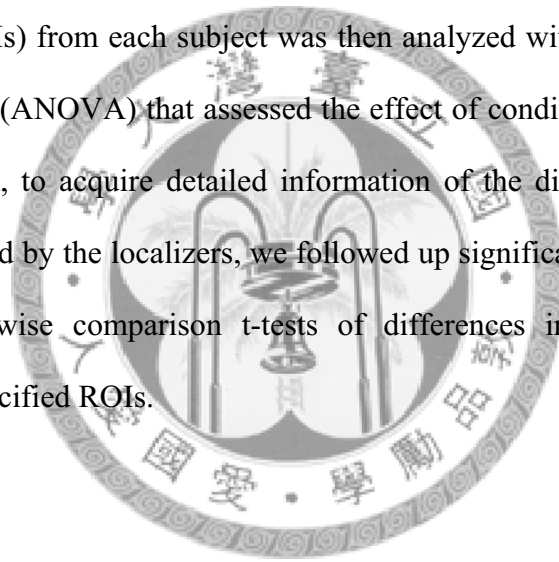
2.34 x 4 mm). The main experiment lasted 225 s (75 images). The first 9 s (3 images) were excluded from further analyses to avoid the start-up transient. Thus, the data analyzed for each scan spanned 216 s (72 images). To correct head-motion artifacts, we used SPM2 (Wellcome Department of Imaging Neuroscience, University College London; <http://www.fil.ion.ucl.ac.uk/spm/>) to realign the acquired EPI images. The realigned images, as well as the anatomic images, were then normalized to a standard template using SPM. The normalized images were fed into the mrVista analysis package (Wandell, et al., 2000) for coregistration, data analysis, and 3D visualization. Statistical analysis of the BOLD activation was based on the spectral correlation between the BOLD activation time series and the experimental sequence (Engel, et al., 1997).



For each participant in each run, the Fourier transform was applied for each BOLD time series. The spectral correlation or coherence for each voxel was computed as the amplitude of the stimulus frequency divided by the summation of amplitude of all frequencies up to the Nyquist limit. The differential activation of a voxel was considered significant if its coherence was greater than .475. This criterion was equivalent to an α -level about 10^{-6} for each individual voxel and Bonferroni corrected α -level, based on the number of gray matter voxels in a flatmap (see Figure 4 & 5), less than .001. The coherence value and phase of each voxel was also used to calculate fitted response amplitude of the time series (i.e., “ β ” in the general linear model) for the group analysis.

The group, or second level, analysis was basically a linear regression of the fitted response amplitude across participants (Penny, Holmes, & Friston, 2003). The second level analysis only considered voxels that were significant at the first level for all

participants. The voxels with significant activation reported in Results had an α -level of .05 for the second level t-test group analysis. For our data, considering cortical voxels, the α -level .05 corresponded to the false discovery rate .07, this is below the “reasonable range” of 0.1-0.2 recommended by Genovese, Lazar, & Nichols (2002) and 0.1 level suggested by Mosig et al., (2001). This criterion was to ensure the activation level surpassed that of the inter-participant variability. To test differential contrasts across different runs, we followed the procedure proposed by Joseph, Partin, and Jones (2002). We first identified regions of interest (ROIs) defined by a specific parametric contrast for individual runs. The signal amplitude of the average waveform over the voxels in these (ROIs) from each subject was then analyzed with a repeated measure analysis of variance (ANOVA) that assessed the effect of conditions for the specified contrast. In addition, to acquire detailed information of the differential contrasts in specific ROIs defined by the localizers, we followed up significant main effects in the ANOVA with pairwise comparison t-tests of differences in activation between conditions in the specified ROIs.



4.2 Results

Psychophysics

Figure 4.2 shows the proportion of correct responses as a function of contrast levels for the real characters (top row, panels a and b) and the non-characters (bottom row, panel c and d) at 1° (left column, panels a and c) and 4° (right column, panels b and d) eccentricity. For all conditions, the proportional correct responses increased with contrast until they reached an asymptotic level. Such data was fit into the function

$$p(x) = \gamma + (1-\gamma) \times \rho \times \Phi(x, \mu, \sigma) \quad (1)$$

where $p(x)$ was the proportional correct response at contrast level x ; γ was the guessing factor for 2AFC and was fixed at .5; ρ determined the asymptotic level of the psychometric function, $\Phi(\mu, \sigma)$ is the Gaussian cumulative distribution function with the location parameter (“mean”) μ and the scale parameter (“standard deviation”) σ . The smoothing curves in Figure 4.2 were fits of this function. The root of the mean squared error (RMSE) of the model fit was about .02 which is close to the mean of the standard error of measurement .04. The values of parameters are shown in Table 4.1.

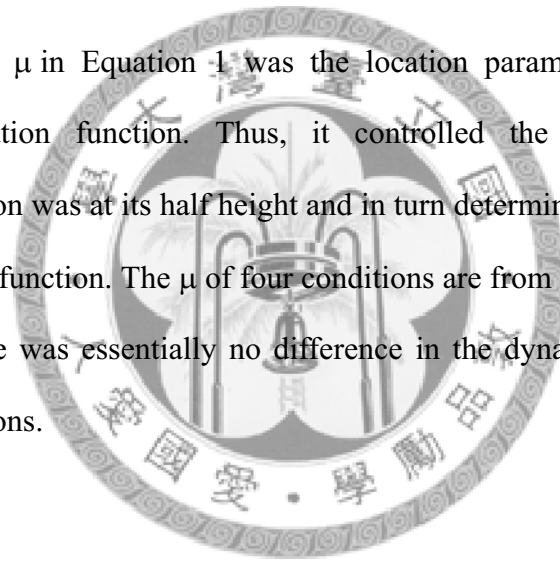
At high contrasts, the proportional correct responses for matching two upright real characters were significantly better than those for matching two inverted real characters at α -level .05 (denoted by the symbol ‘*’ in Figure 4.2a and Figure 4.2b). In the parafovea viewing conditions, the psychometric function for upright real characters was asymptotic to the percentage correct level of 82%, while the function for inverted characters was 67%. In the peripheral viewing conditions, the psychometric function for upright real characters was asymptotic to the percentage correct level of 80% while the function for inverted characters was 64%. That is, in both parafoveal and peripheral viewing conditions, at high contrasts, turning the real characters upside down reduced the percentage correct level by about 15%.

Figure 4.2c and Figure 4.2d show the psychometric functions for matching non-characters. There was no difference between the proportional correct responses for matching upright and matching inverted non-characters in either the parafovea or in the periphery. The psychometric function was asymptotic to the level 77% and 70% for upright and inverted characters respectively in the parafovea, and was asymptotic to 68% and 64% for upright and inverted characters respectively in the periphery. The

decline in performance was about 6% and showed no statistical significance. Thus, there is no inversion effect for non-characters in either the parafovea or the periphery.

Figure 4.3 replots the data for upright characters, to compare the matching performances for real and non-characters. With the exception of the lowest test contrast where the performance is always at the chance level, the matching performance for real characters was better than that for non-characters at all contrasts and the magnitude of the difference is statistically significant (α -level = .05) at high contrasts. The difference is even more pronounced in the peripheral viewing condition.

The parameter μ in Equation 1 was the location parameter of the Gaussian cumulative distribution function. Thus, it controlled the position where the psychometric function was at its half height and in turn determined the dynamic range of the psychometric function. The μ of four conditions are from -16.648 to -19.126 dB contrast level. There was essentially no difference in the dynamic range for all the psychometric functions.



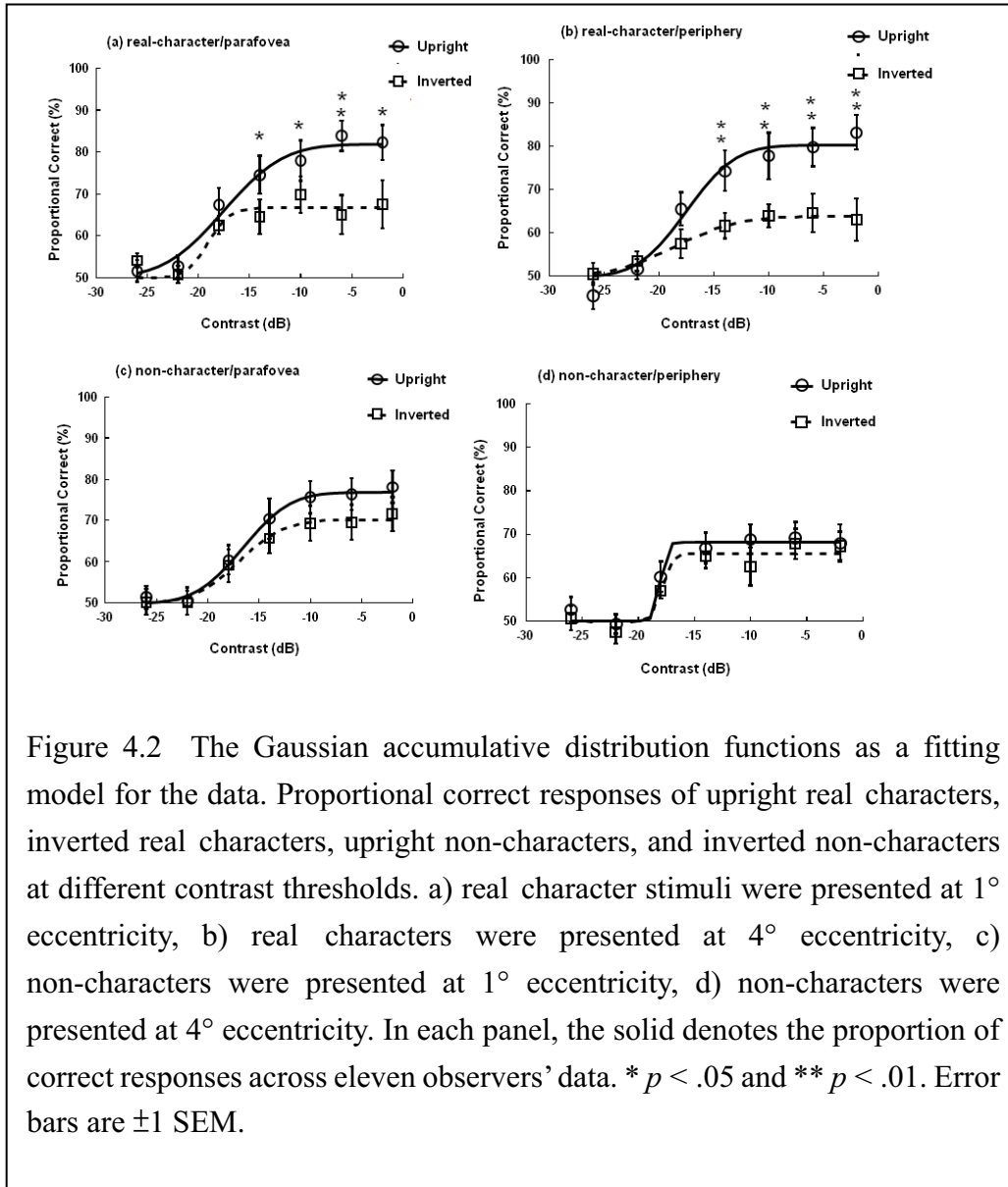


Figure 4.2 The Gaussian accumulative distribution functions as a fitting model for the data. Proportional correct responses of upright real characters, inverted real characters, upright non-characters, and inverted non-characters at different contrast thresholds. a) real character stimuli were presented at 1° eccentricity, b) real characters were presented at 4° eccentricity, c) non-characters were presented at 1° eccentricity, d) non-characters were presented at 4° eccentricity. In each panel, the solid denotes the proportion of correct responses across eleven observers' data. * $p < .05$ and ** $p < .01$. Error bars are ± 1 SEM.

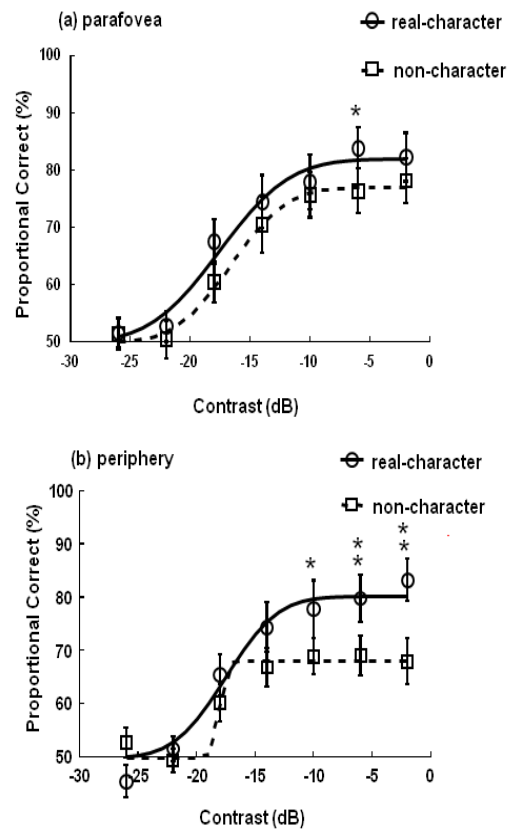


Figure 4.3 Proportional correct responses of upright real and non-characters at different contrast thresholds. a) real characters and non-characters were presented at a 1° eccentricity, b) real characters and non-characters were presented at a 4° eccentricity. * $p < .05$ and ** $p < .01$.

Table 4.1 The resulting model fits to combined data set of eleven observers. The parameter μ is the mean of the contrast, σ and ρ is the scale parameter, and RMSE indicates the goodness-of-fit.

Parafovea	μ	σ	ρ	RMSE
Upright real-character	-17.5887	4.5907	32.0973	0.0189
Inverted real-character	-19.1264	1.7208	16.7225	0.0221
Upright non-character	-16.6484	3.6514	27.0651	0.0095
Inverted non-character	-17.0390	3.5190	20.2507	0.0099
Periphery				
Upright real-character	-17.5505	3.4870	30.2563	0.0248
Inverted real-character	-18.5808	4.5234	13.8820	0.0049
Upright non-character	-18.0782	0.4955	18.1800	0.0126
Inverted non-character	-17.9132	0.7077	15.6250	0.0185

Behavioral performances in the fMRI experiment

The 1-back matching task was used here to keep participants' attention as it was also used in some fMRI studies on face inversion effects (e.g., C. C. Chen, et al., 2007; Yovel & Kanwisher, 2005). Nevertheless, we did record the performance of 11 (out of 13) participants. While the purpose of this task was simply attention control and the number of trials collected during scanning was small, we did find behavioral inversion effects in the 1-back task. The proportional correct response for upright real characters was higher than that for inverted real characters (96% vs. 93%, pairwise $t(10) = 3.187$, $p < .05$). However, there was no difference in performance for upright and inverted non-characters (95% vs. 94%, pairwise $t(10) = 1.44$, $p > .05$). This result was consistent with the main psychophysical experiment reported above. Note that, since the averaged proportional correct responses were all greater than 93%, we did not separate the cortical activation for correct and incorrect trials in our fMRI data analysis.

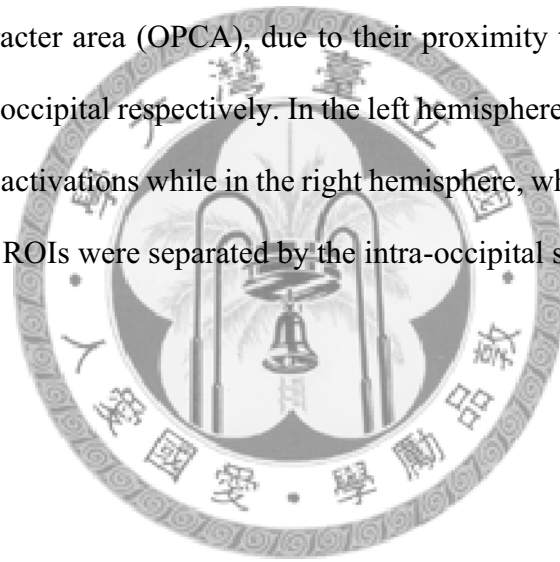
Cortical activation to upright characters

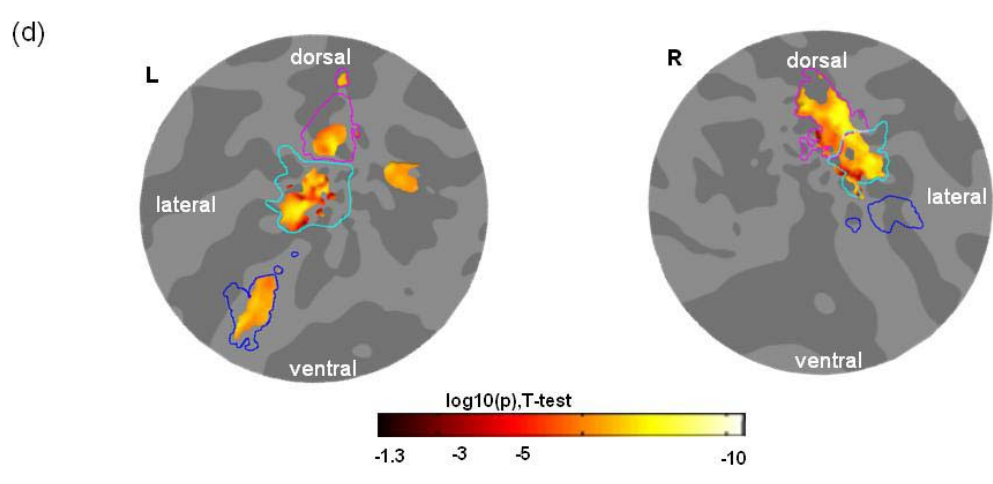
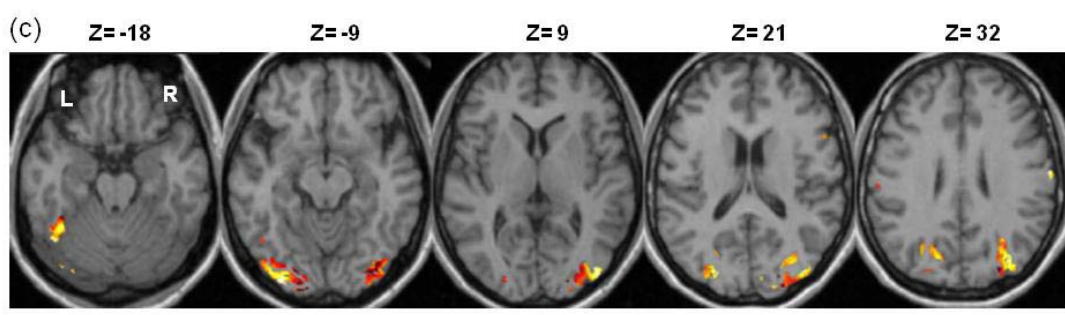
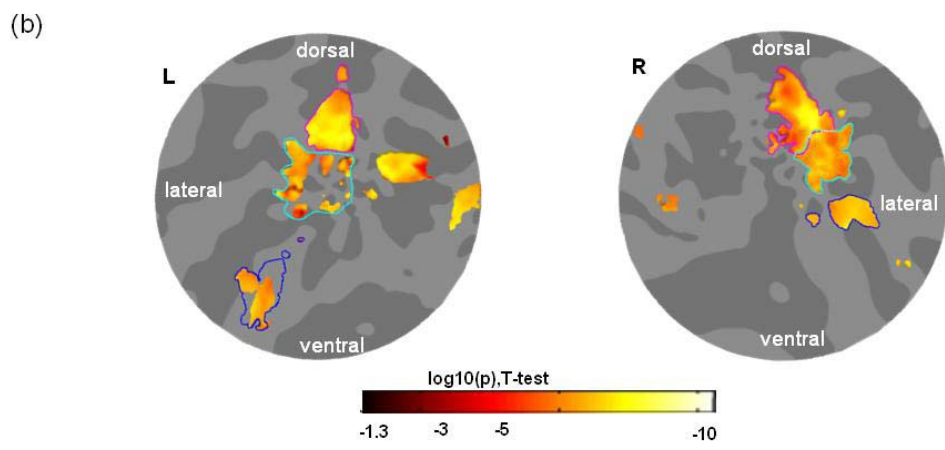
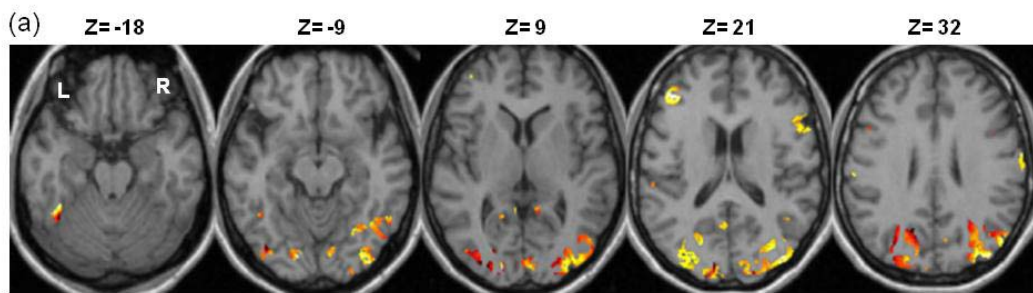
Figure 4.4a shows the differential activation for contrasting real characters with their own scrambled versions in axial slices from a group analysis. The pseudo-colored patches in the figures denote voxels showing a significant activation change between real characters and their scrambled versions at α -level .05. The majority of activated voxels were located in the occipitotemporal regions, including the fusiform gyrus, the middle and the superior occipital gyri (MOG and SOG, respectively) in both hemispheres and the inferior occipital gyrus (IOG) in the left hemisphere. The activated areas on the lateral surface overlapped with the lateral occipital complex

(LOC) reported in the literature (Kourtzi & Kanwisher, 2001). Small patches of activation were also observed in the inferior frontal gyrus (IFG), the left middle temporal gyrus (MTG) and the right postcentral gyrus. Table 4.2 lists the Talairach coordinates (Talairach & Tournoux, 1988) of the major activated areas. The left fusiform showed a more extensive activation than the right fusiform ($\chi^2(1)=428.96, p < .05$). This is consistent with what was reported in the literature regarding the visual word forms in English (L. Cohen, et al., 2000; L. Cohen, et al., 2002). On the other hand, the right dorsal occipital cortex showed more extensive activations than the left ($\chi^2(1)=1153.42, p < .05$). This is consistent with the studies which used logograph words as stimuli (Tan, Feng, et al., 2001; Tan, Liu, et al., 2001). For a better understanding of the spatial relationships between the activated areas, we replotted the activation pattern on a flatmap (in Figure 4.4b). The flatmaps in this paper were 100 mm radius disks centered at a point near the occipital poles on inflated cortical surfaces. The gray shading denotes the gyral (light) and sulcal (dark) layouts. The criteria for choosing ROIs will be discussed below.

Figure 4.4c shows the group-averaged activation map for contrasting non-characters with scrambled versions of characters in axial slices. The pseudo-colored patches denote voxels showing significant activation change differences between non-characters and their scrambled versions at α -level .05. The majority of activated voxels were also located in the occipitotemporal regions, including the left fusiform gyrus, the bilateral MOG, SOG and the right cuneus. Small patches of activation were also observed in the IFG and right postcentral gyrus. The activation pattern of non-characters is also shown in Figure 4.4d. Talairach coordinates of the major activated areas are also given in Table 4.2.

For the convenience of discussion, we defined three ROIs in each hemisphere from the activation maps for the real and non-characters conditions. The colored contours in Figure 4.4 show the boundary of these ROIs: the fusiform character area (FCA, blue contours), the lateral occipital character area (LOCA, cyan contours) and the occipitoparietal character area (OPCA, magenta contours). A voxel is a part of an ROI if it shows significant differential activation either in the real or in the non-character conditions. The FCAs were on a cluster of activation in the middle fusiform gyrus. The character selective activations in the dorsal occipital cortex were separated into two areas of interest: the lateral character area (LOCA) and the occipitoparietal character area (OPCA), due to their proximity to the lateral occipital complex and kinetic occipital respectively. In the left hemisphere, these two ROIs were separated clusters of activations while in the right hemisphere, where there was no clear boundary, these two ROIs were separated by the intra-occipital sulcus.





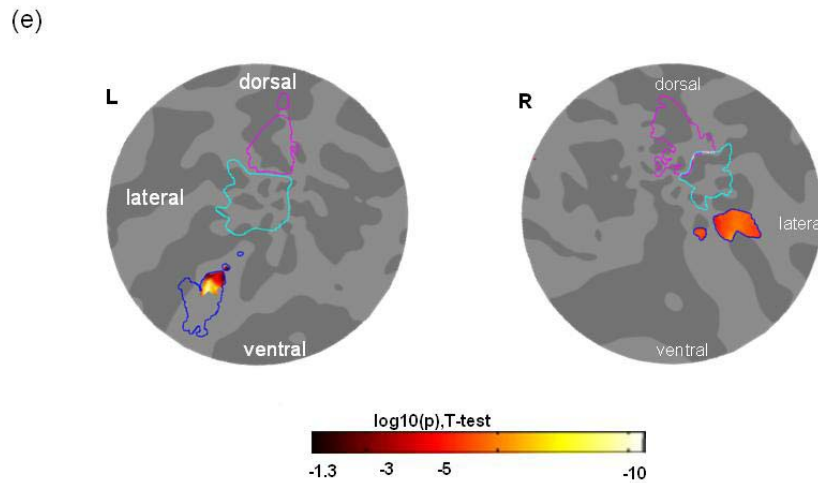


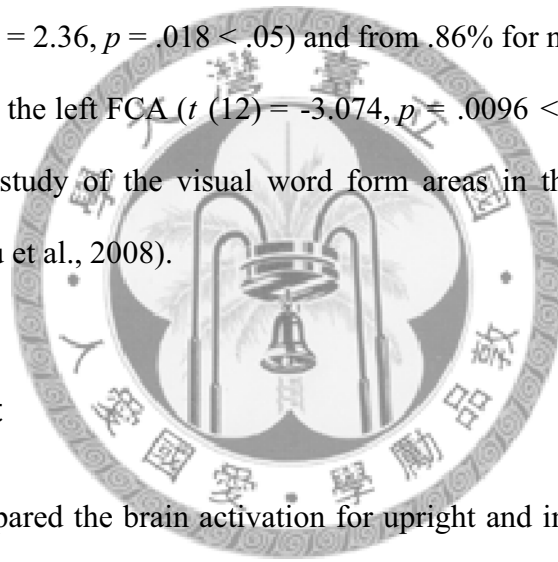
Figure 4.4 (a) Activation maps are shown in the standardized brain slices for the real characters versus their scrambled versions. Orange color denotes the average activation pattern. (b) The real characters activation depicted in flatmaps centered on the occipital poles. (c) Activation maps are shown in the standardized brain slices for the non-characters versus their scrambled versions. (d) The non-characters activation depicted in flatmaps centered on the occipital poles. (e) Differential activation between real-characters and non-characters in flatmaps. The blue border denotes the extent of the activated areas in the FCA. The cyan border denotes the extent of the activated areas in the LOCA. The magenta border denotes the extent of the activated areas in the OPCA. Gyral (light) and sulcal (dark) denoted by shading. Yellow-orange coloration codes the negative log base 10 of the significance value of a 2-side t -test (threshold at $p < .05$). L: left hemisphere; R: right hemisphere.

Table 4.2 The brain areas showing differential between characters and scrambles

Brain Regions	Hemisphere	Talairach Coordinate			Volume (mm ³)
		X	Y	Z	
<i>real-characters</i>					
Fusiform gyrus	left	-41	-58	-16	1816
	right	36	-64	-9	764
MOG & IOG	left	-31	-90	20	1015
		-39	-84	0	1435
	right	38	-86	-5	5473
SOG	left	-34	-88	24	2292
	right	32	-76	28	1346
<i>non-characters</i>					
Fusiform gyrus	left	-43	-58	-16	1597
MOG	left	-37	-85	-8	3232
	right	38	-87	-4	4768
SOG	left	-29	-88	20	351

Differential activations between real and non-characters

We then applied a conjunction analysis (Joseph, et al., 2002) to study the brain areas showing differential activation to real and non-characters (see Method). Figure 4.4e shows the areas with significant differential activation between real and non-characters on flatmaps with ROIs. The FCA of both hemispheres showed a significant activation difference between the real characters and non-characters ($F_{12,24} = 6.16, p = .03 < .05$ for the left and $F_{12,24} = 4.72, p = .05$ for the right). The percentage of amplitude change dropped from .62% for real characters to .41% for non-characters in the right FCA ($t(12) = 2.36, p = .018 < .05$) and from .86% for non-characters to .66% for real characters in the left FCA ($t(12) = -3.074, p = .0096 < .05$). This result was consistent with the study of the visual word form areas in the fusiform gyrus for reading Chinese (Liu et al., 2008).



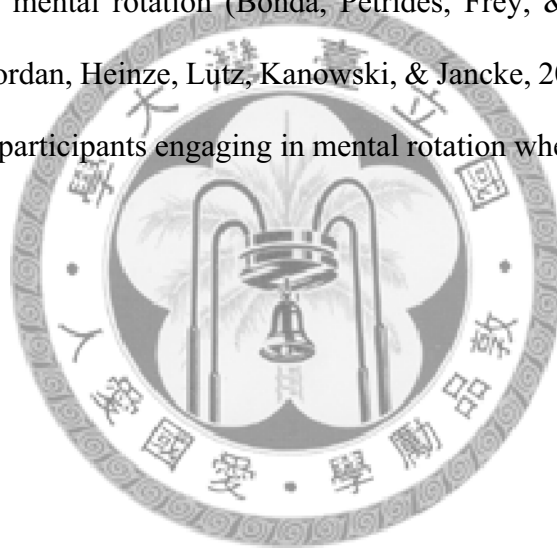
Inversion effect

We then compared the brain activation for upright and inverted characters. As shown in Figure 4.5a, contrasted with their inverted versions, the upright real characters showed robust differential activation in the left FCA and the LOCA defined above. Table 4.3 lists the Talairach coordinates of the activated areas. Figure 4.5b shows that contrasted with their inverted version, non-characters showed no differential activations greater than the noise level in any voxels.

The amplitudes of the inversion effect in different areas are shown in Figure 4.6. With the conjunction analysis, we found that the inversion effect for the real characters in the left FCA was significantly greater than that for the non-characters ($F_{12,24} = 18.17,$

$p = .001 < .05$). The percentage of amplitude change dropped from .83% for real characters to .34% for non-characters ($t(12) = 4.49, p = .00037 < .05$). The bilateral LOCA also showed an inversion effect for the real characters greater than that for the non-characters ($F_{12,24} = 5.48, p = .037 < .05$ for the left hemisphere; $F_{12,24} = 5.50, p = .037 < .05$ for the right hemisphere). The percentage of amplitude change dropped from .96% to .54% ($t(12) = 2.57, p = .012 < .05$) in the left LOCA, and from .83% to .47% ($t(12) = 2.50, p = .014 < .05$) in the right LOCA.

Note that we found no activation in the precuneus, which was considered to be an area associated with mental rotation (Bonda, Petrides, Frey, & Evans, 1995; M. S. Cohen et al., 1996; Jordan, Heinze, Lutz, Kanowski, & Jancke, 2001). Thus, we did not find evidence of our participants engaging in mental rotation when processing inverted characters.



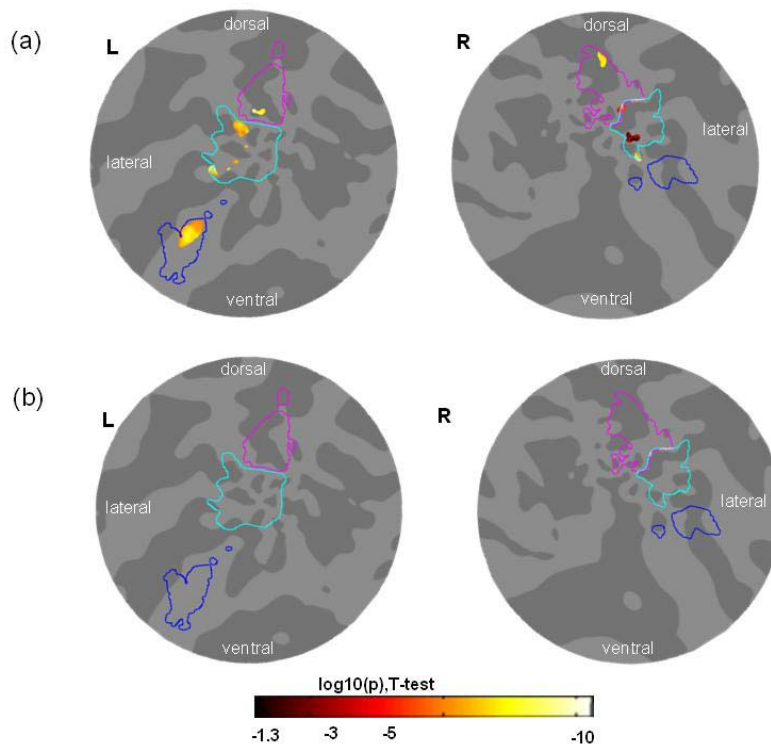


Figure 4.6 (a) Differential activation between upright and inverted real characters and (b) differential activation between upright and inverted non-characters depicted in flatmaps centered on the occipital poles. Blue borders: FCA; cyan borders: LOCA; magenta borders: OPCA. Activation colors as in Figure 4.4.

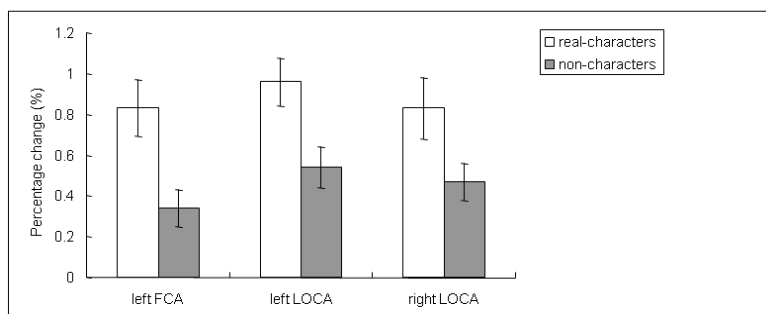


Figure 4.5 Percentage change of activation between the upright and inverted real and the non-characters in FCA and LOCA. The error bars denote the standard error calculated across participants.

Table 4.3 The brain areas showing inversion effects of real characters

Brain Regions	Hemisphere	Talairach			Volume (mm ³)
		Coordinate			
		X	Y	Z	
Fusiform gyrus	left	-44	-62	-15	641
IOG	left	-34	-89	-3	582
	right	41	-87	-4	269

4.3 Discussion

In the psychophysical experiment, we showed that the proportional correct responses for matching upright real characters were significantly better than those for matching inverted real characters. Such an inversion effect was not observed in non-characters. The matching performance for upright real characters was better than that for upright non-characters. In addition, the psychometric functions showed the same dynamic range in luminance contrast for all character types. With the fMRI experiment, we found that part of the left fusiform gyrus and a small area in the bilateral lateral occipital regions had a significant differential activation between upright and inverted real characters. The fusiform gyri in both hemispheres also showed differential activation between upright real and non-characters. The occipitoparietal regions

showed the character selective activation when characters were compared with scrambled lines, but was indifferent to any manipulations of characters used in this study.

Behavioral evidence for inversion effects

The proportional correct responses for matching real characters reduced significantly when the characters were turned upside-down. This inversion effect was reported only with well-practiced visual objects, such as faces (Carey & Diamond, 1977; Leder & Bruce, 2000; Rhodes, et al., 1993; Yin, 1969). The inversion effect, as suggested by Carey and Diamond (1977), reveals the spatial configuration processing in the visual analysis of those objects. After all, the relative spatial relations among image elements are turned upside down in the inverted images. Hence, the response of a mechanism that is specialized for the spatial configurations of image elements would be reduced by the inversion. Thus, our result may suggest that a spatial configuration process is also involved in character processing.

We also showed that the matching performances for real characters were significantly better than those for the non-characters. In addition, the proportional correct responses for upright non-characters were similar to those for the inverted real and non-characters. As discussed above, a non-character was constructed by swapping the left and right parts of a real character. Hence, the non-characters disrupted the global configuration while keeping local configuration intact. Thus, our result suggests that the function of the character processors in the visual system is to analyze the global configuration among character components and not local information.

The psychometric functions showed the same dynamic range in luminance contrast for all character types. The location parameters of all the fitted psychometric functions were almost identical. This result may suggest that the early visual mechanisms were not decisive factors in our character matching tasks. The early visual mechanisms are sensitive to the local orientations and contrast in an image (Campbell & Robson, 1964; Hubel & Wiesel, 1962, 1969). The local orientations of different types of characters varied substantially in our stimuli. And yet, there is no contrast effect across character types. Therefore, the mechanisms underlying character matching tasks must be able to take the whole character, rather than local features, into account.

Strong inversion effects for real characters were found both in the parafovea and in the periphery. This result was consistent with the face inversion effect which was also found both in the fovea and in the periphery (McKone, et al., 2007). However, it was reported that, after a few hours of training, the inversion effect for trained objects can emerge in the fovea but not in the periphery (McKone, Martini, & Naykayama, 2003). Such eccentricity difference in the inversion effect has been used as an argument for the uniqueness of face processing (McKone, et al., 2003). Our result suggests that the eccentricity invariant inversion effect is not unique to faces. After all, visual word forms, similar to faces, are well-practiced visual stimuli. Perhaps the mechanisms for visual word form processing shared more common properties with those for face processing than those for object processing.

The role of the fusiform gyrus in character processing

It has been reported that part of the left fusiform gyrus specifically responds better to words than nonsense letter strings (non-words) and was named the visual word form area (VWFA) by some authors (e.g., Cohen et al., 2000, 2002). On the other hand, it is also suggested that the area designated as VWFA may also respond to other categories of visual objects (e.g., Price & Devlin, 2003). Similarly, another part of the fusiform gyrus was reported to be specific for face processing, termed the fusiform face area or FFA (Kanwisher, et al., 1997). Again, there is an argument that the FFA is also responsible for other types of visual objects (Gauthier, Behrmann, & Tarr, 1999; Haxby et al., 2001). In particular, it was suggested, the FFA may be responsible for expertise in perceiving objects (Gauthier & Tarr, 1997; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Since both word and face recognitions are well-practiced visual tasks, it is difficult to separate category dependent and familiarity dependent responses from faces and words in the FFA.

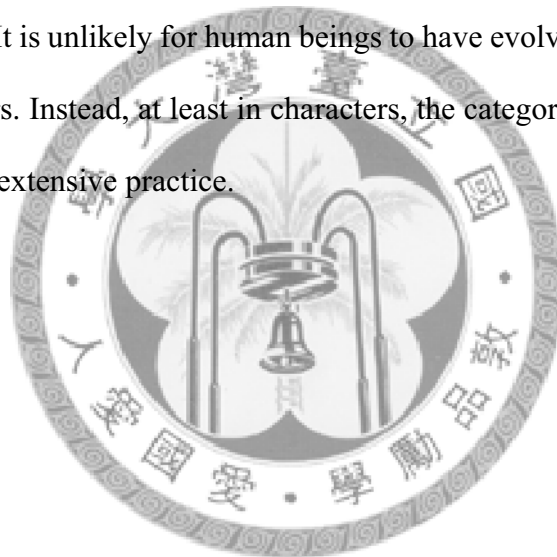
In this study, we found that the fusiform gyrus showed greater activation for the upright real characters than the inverted ones. Previous studies also reported that the fusiform gyrus responded better to upright faces than inverted faces (C. C. Chen, et al., 2007; Yovel & Kanwisher, 2004, 2005). The inversion effect we found in the study is consistent with that reported in previous studies on face inversion effect in the FFA. One possible interpretation is that the neurons in the fusiform are insensitive to inverted stimuli.

At the behavioral level, the inversion effect to an object category seems to relate to the experience of an observer to that object category. If the function of the fusiform is

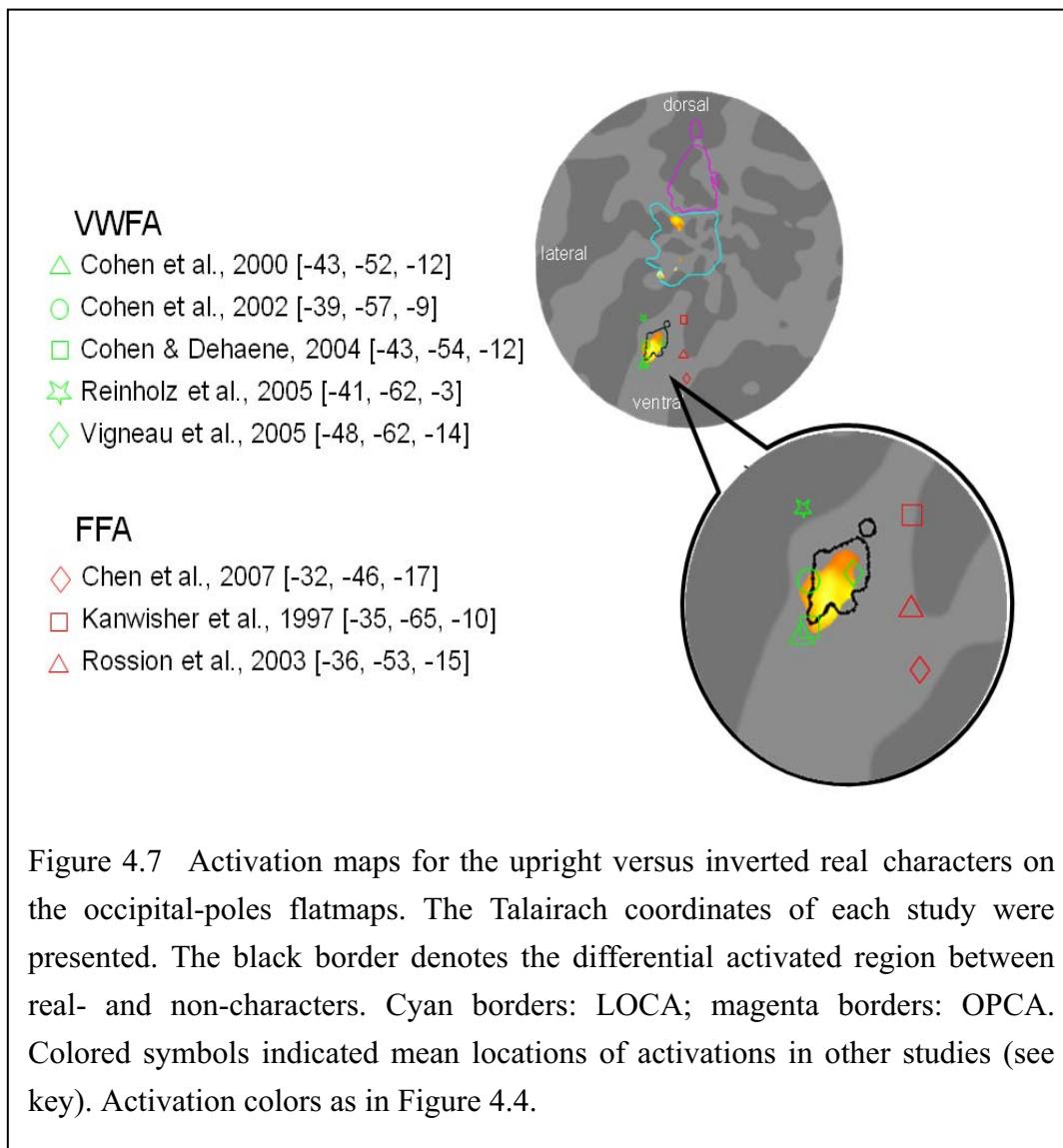
for category-specific processing, we would expect the character inversion effect to be limited to character-selective areas. On the other hand, if the fusiform gyrus is related to expertise and is not category specific, we should expect the area showing an inversion effect to be widely spread.

While both words and faces can increase activation in the fusiform, the peak activations occur in different parts of the left fusiform gyrus. The green and red symbols in Figure 4.7 denote the locations of peak activation for words (L. Cohen & Dehaene, 2004; L. Cohen, et al., 2000; L. Cohen, et al., 2002; Reinholz & Pollmann, 2005; Vigneau, Jobard, Mazoyer, & Tzourio-Mazoyer, 2005) and for faces (C. C. Chen, et al., 2007; Kanwisher, et al., 1997; Rossion et al., 2003) respectively. The peak locations were computed from the reported Talairach coordinates in those papers. The word activations tend to be located in regions that are lateral to the regions for face activations. The black contour in Figure 4.7 indicates the region showing significant differential activation between the real and the non-characters as shown in Figure 4.4e. That is, this region is an equivalent of VWFA, which was defined as the brain region showing differential activations between words and non-words, for Chinese characters. This region is similar to the VWFA reported in the literature. This suggests a similarity in the visual processing of characters in Asian languages and word processing in alphabetic languages. The colored patch denotes the area showing the inversion effect as in Figure 5. The area showing the inversion effect overlapped with the VWFA but not with face-selective regions. This result implies that, (1) the function of VWFA is to analyze global spatial configurations among character components (Starrfelt & Gerlach, 2007) and (2) the character inversion effect is character-specific processing. The former may be consistent with the report that

English-speaking pure alexia patients may have an impairment reading words, but have normal letter identification performances (Farah & Wallace, 1991; Sekuler & Behrmann, 1996). It suggests that expertise alone cannot explain the function of the fusiform gyrus. Instead, the function of the left fusiform gyrus shows a certain degree of categorical specification. Note that we do not suggest that expertise plays no role in the function of the fusiform gyrus. After all, the fusiform gyrus showed strong differential activation between the familiar and the novel after training (Baker et al., 2007; Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier, Tarr, et al., 1999). In addition, written languages, being words or characters, were only invented less than ten thousand years ago. It is unlikely for human beings to have evolved a brain specifically for reading characters. Instead, at least in characters, the categorical specification may be acquired through extensive practice.



The right fusiform gyrus showed a difference in activation only between real and non-characters. However, the inversion effect is absent in this region. This may imply that the right fusiform gyrus is not sensitive to the orientation of character components.



The dorsal activation

A small portion of the LOCA showed an inversion effect while the OPCA did not. None of these areas showed a significant difference between the activations to real and non-characters. The lateral occipital regions are known to be responsive to visual

objects (Kanwisher, Chun, McDermott, & Ledden, 1996; Kourtzi & Kanwisher, 2000, 2001). Hence, it is surprising that these areas responded differently to characters and scrambled line segments. The LOCA responded differently to upright and inverted characters, which were different in both global and local configurations. However, this region was indifferent to real and non-characters, which were different only in global configurations and it showed no differential activation between upright and inverted non-characters, which were different in local configurations. Neither carried any familiar global configuration. That is, the differential activation in the LOCA requires a difference in both the local and global configurations. This may be consistent with the notion of the involvement of the lateral occipital in integrating local fragments of an object for the global analysis in the brain areas further down stream. (Hayworth & Biederman, 2006; Ostwald, et al., 2008).

The OPCA was indifferent to all character manipulations. Hence, it may be that the OPCA responds only to local features, such as corners and junctions that were changed when we scrambled the characters. Previously, studies on character recognition showed that the occipitoparietal region was involved in the processing of the spatial arrangement of strokes or character components (Fu, Chen, Smith, Iversen, & Matthews, 2002; Kuo, et al., 2001; Tan, Feng, et al., 2001; Tan, Liu, et al., 2001).

The real characters produced an extensive activation in the right dorsal occipitoparietal region. Such activation was not often reported in the brain imaging studies of words in alphabetic languages (e.g., Fiez & Petersen, 1998). Instead, the occipitoparietal activation to visual word forms in alphabetic languages seemed to be confined to the left hemisphere. On the other hand, the extensive right occipitoparietal activations, as we report here, were often reported in studies with words in logograph

languages (Tan, Feng, et al., 2001; Tan, Liu, et al., 2001). Such right hemisphere activation was often attributed to the requirement of more intensive spatial analysis of the written words in logograph languages (Tan, Feng, et al., 2001; Tan, Liu, et al., 2001).

Hierarchical processing in the occipitotemporal cortex

From the above discussion of our fMRI results, we conclude that from dorsal to ventral, the visual system may analyze characters in a hierarchical way. The occipitoparietal regions may analyze the local features of an object regardless of its familiarity. Then, the lateral occipital regions may play an intermediate role in integrating the local information in an object. Finally, the fusiform gyrus plays a critical role in analyzing global configurations in a character. This is consistent with the notion that the human visual cortex analyzes the image of an object in a hierarchical way. That is, while the local features in an image are analyzed in the early visual cortex, the information processing is becoming more global and more complicated as the information flows to brain areas further downstream.

Such hierarchical processing was also found in studies with other visual stimuli. Ostwald et al. (2008) showed that the activations in the early visual areas were influenced by the local features, such as dipole orientation, of a Glass pattern (Glass, 1969), which those in the lateral occipitotemporal regions were determined by the global forms. Similarly, Kourtzi and Huberle, (2005) also found the early visual areas were sensitive to the local orientations of contour elements, whereas the lateral occipital complex was sensitive to the global form of the contours.

For a more complicate visual stimulus, Chen et al. (2007) showed that the activations in the FFA were influenced by the global configuration of a face, such as symmetry, while the activations in the dorsal occipital cortex were only influenced by local features. Yovel and Kanwisher (2005) showed that the face inversion effect was mainly observed in the FFA and the effect is not very robust in the occipital cortex. Taylor, Wiggett, and Downing (2007) showed that the activation in the occipital cortex is selective to parts of a body whereas the activation in the fusiform is selective to the configuration of body parts. These studies drew similar conclusions to ours with regard to visual word form.

However, studies with Navon figures (Navon, 1977) showed a different picture. A Navon figure is a large letter with strokes composed of smaller letters. Han et al. (2002) showed that the medial occipital was activated when participants gave responses based on the larger letter (global feature) and the parieto-occipital sulcus, when based on the smaller letter (local feature). With a similar task, Han, Jiang, and Gu (2004) showed that the temporal cortex was activated when the participants were asked to attend to the larger letter, and the superior parietal cortex, when they were asked to attend to the smaller letter. Despite the inconsistency in the Navon's figure literature itself, there is one major difference between Navon's figure studies and ours. Similar to the Glass pattern contour integration, face and body part studies discussed above, our study focused on the visual analysis of local features and global configurations while the Navon's figure studies focused on the local and global attention strategy. Hence, these two types of studies may tag into completely different neural mechanisms. This notion is consistent with the finding of Davidoff, Fonteneau, and Fagot (2008) who found that

people from a remote culture had global representations only for faces but not for Navon's figures.

4.4 Summary

Taken together, we used psychophysical and functional MRI methods to investigate the character processing. The behavioral evidence showed that performance for matching upright real characters was significantly better than that for matching inverted real characters. Such an inversion effect was absent for non-characters. We also found that the matching performance for upright real characters was better than that for upright non-characters. In addition, the psychometric functions showed the same dynamic range in luminance contrast for all character types. These results suggested that the recognition of real characters involved spatial configuration processing. In the fMRI evidence, the left fusiform gyrus and a small area in the bilateral lateral occipital regions showed a significant differential activation between upright and inverted real characters. The bilateral fusiform gyri also show differential activations between upright real and non-characters. The occipitoparietal regions showed character selective activation when compared with scrambled lines. It suggests that the visual system may analyze the spatial configurations of a visual word form in a hierarchical processing.

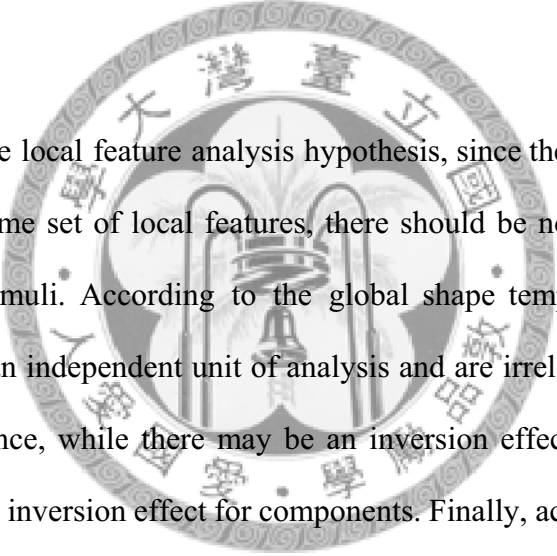


Chapter 5 : The inversion effect in visual word form processing: the component effect

In Chapter 4, we found that the visual system may analyze the spatial relationship, or configuration, among character components. The participants' performances in matching upright real characters were significantly better than in matching inverted real characters. This inversion effect was not observed in non-characters, suggesting that recognizing a visual word depends not only on the shapes of individual features, but also on the relationships among them. The left FG plays a critical role in processing the spatial configuration of a visual word. On the other hand, the early visual cortex is involved in analyzing the local features in a visual word.

In this chapter, we further investigate the relationship between the local and the global configuration in characters, by comparing the inversion effect of a character and its components. In Chapter 1, while we found that the contrast thresholds for discriminating between real and pseudo-characters, non-characters were significantly higher than that for *Jiagu* and scrambled characters. Because real, pseudo-, and non-characters are constructed by two familiar components, the visual system may analyze the global shape of a character or the components in a character. In order to understand the role of components of a character in processing visual word forms, we use psychophysics and fMRI to compare the inversion effect of a character and its components

We used eight types of orthographic stimuli, including real characters, non-characters, lexical components, non-lexical components of a character, and their inverted versions, to distinguish the local configurations, or the spatial relations among strokes, from the global configurations, or the spatial relations among components. A certain left-right configurative character has one component that can be used as an independent character in its own right (e.g. the right side “亊” of a composite character “對”) while the other component can only be used in a composite character (e.g. the left side “耑” of a composite character “對”). For the purposes of this discussion, we referred to the former as the lexical component and the latter as the non-lexical component.



According to the local feature analysis hypothesis, since the upright and inverted words contain the same set of local features, there should be no inversion effect for visual word form stimuli. According to the global shape template hypothesis, the components are not an independent unit of analysis and are irrelevant for visual word form processing. Hence, while there may be an inversion effect for real characters, there would not be an inversion effect for components. Finally, according to the spatial configuration analysis hypothesis, the visual word form processing may involve multiple stages in the visual system: called the local and the global stages in this study. The local stage analyzes the spatial relationship among strokes in a character component while the global stage analyzes the spatial relationship among components in a character. The local stage is to create character components from strokes while the global stage combines components to form a character. The holistic processing of characters may start from either one of the stages. If the holistic processing starts from the local stage, we should be able to observe the inversion effect for both the lexical and

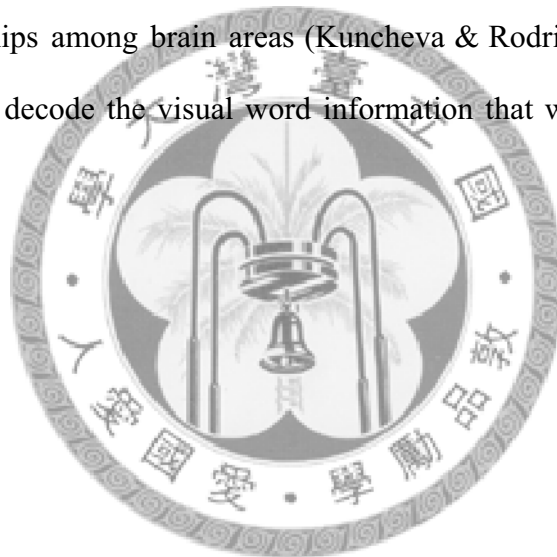
non-lexical components. However, if the holistic processing only starts from the global stage, then we should not be able to observe the inversion effect for lexical and non-lexical components.

Furthermore, we used multivariate analysis with fMRI to investigate whether cortical areas are involved in the analysis of the features of a character, character specific configurations, or other holistic properties of a character. More specifically, we intended to determine what exact computations are conducted in the ventral occipital areas, and what kinds of representations it extracts from visually presented words.

The univariate statistical methods, such as the general linear model (Friston et al., 1995), are widely employed in fMRI studies. In these methods, the BOLD response of each voxel is independently analyzed. This analysis focuses on finding locations that show a significantly different response between two or more experimental conditions (Friston, et al., 1995). However, this individual voxel based analysis may not be sufficient to explore the relationship between human behavior and brain function. First, this analysis assumes that activation of each voxel is independent from each other. However, the activation of one voxel is correlated with that of neighboring voxels. Second, this univariate method assumes that a brain area responds to only one function in the experiment. It cannot dissociate the multiple functions carried by the same area (Kriegeskorte, Goebel, & Bandettini, 2006).

In this study, we used a multivariate method, the support vector machine (e.g., Cortes & Vapnik, 1995) with fMRI to classify the character information in the visual cortex. Unlike conventional fMRI analysis, the multivariate method does not spatially average responses across activated voxels (Haynes & Rees, 2006; Norman, Polyn,

Detre, & Haxby, 2006). Instead of voxel-by-voxel analysis, researchers use powerful pattern classification algorithms to classify the distributed activity patterns across multiple voxels. Furthermore, multivoxel activation patterns allow researchers to simultaneously compare the responses of stimuli across different conditions by the pattern of activation. This analysis was named “multivoxel pattern analysis” (Haynes & Rees, 2006; Norman, et al., 2006). The multivoxel pattern analysis has been applied to the object recognition (Cox & Savoy, 2003; Haxby, et al., 2001; LaConte, Strother, Cherkassky, Anderson, & Hu, 2005), orientation, (Kamitani & Tong, 2005) and attention (Esterman, Chiu, Tamber-Rosenau, & Yantis, 2009), to characterize the functional relationships among brain areas (Kuncheva & Rodriguez, 2010). We first used this method to decode the visual word information that was represented in the visual cortex.



5.1 Methods

Psychophysics

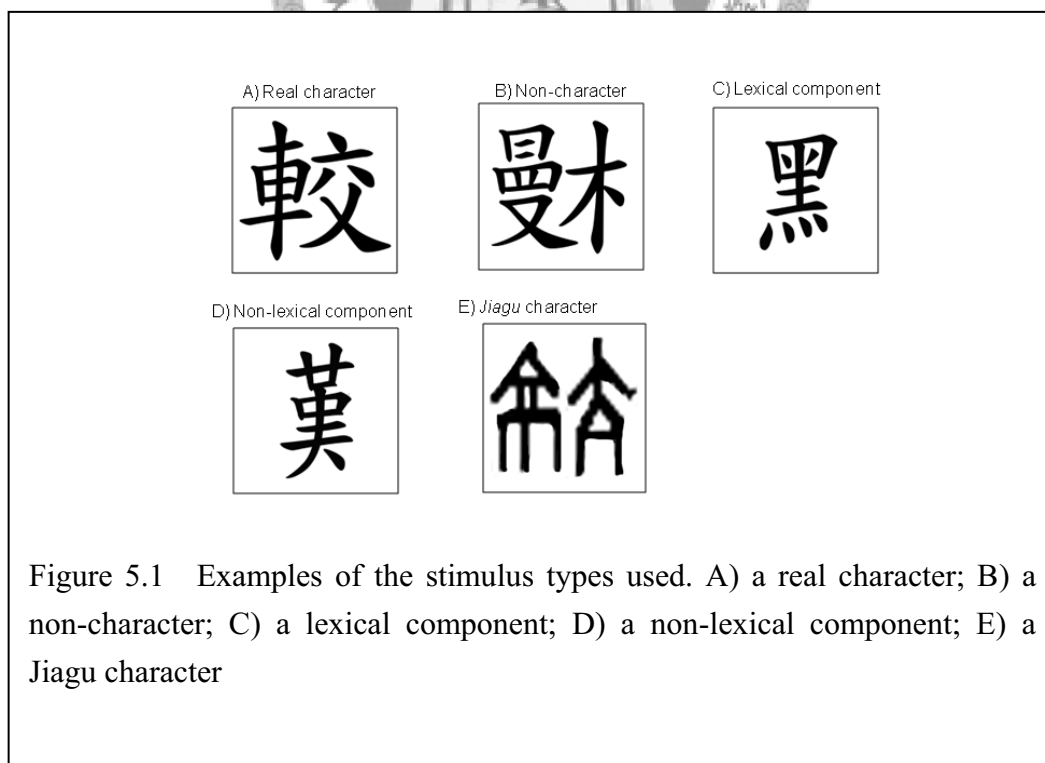
Observers Eleven right-handed observers (four males, seven females) between the ages of 22 and 28 years old participated in this study, receiving financial remuneration for their participation. All participants read Chinese fluently and naïve as to the purpose of the experiment.

Apparatus The stimuli were presented on a monitor controlled by a RADEON 9800 XT video card on a PC. The monitor resolution was 1280 (H) x 1024 (V). The monitor input–output intensity function was measured with a light mouse photometer (Tyler & McBride, 1997). This information allowed us to compute linear lookup table

settings in order to linearize the output. The mean luminance of the monitor was 30 cd/m². At a viewing distance of 54 cm, each pixel on the monitor had a size of .03 degrees in visual angle.

Stimuli Ten types of stimuli were involved in this experiment. The real characters, non-characters, and *Jiagu* characters used in this experiment were the same as those in Chapter 1. Figure 5.1 shows examples of stimuli used in this study. Both lexical and non-lexical components are parts of a composite character. However, the lexical component can be used as an independent character in its own right while a non-lexical component only exists in a composite character. The inverted stimuli were simply the upside-down versions of their upright counterparts. There were 60 stimuli in each type of characters. All the character stimuli were listed in Appendix 3.

The stimuli were presented at 1 degree away from the central fixation along the



horizontal meridian. The stimuli size was 1.6 deg and the contrast level of the stimuli ranged from -26 to -2 dB in 4 dB step. The unit dB is 20 times log base 10 of the linear contrast.

Procedure. The procedure was the same as the one described in Chapter 4. In each trial, two characters of the same type and orientation were presented simultaneously for a duration of 100 ms. One was on the left and the other was on the right of the central fixation. The stimuli were presented at 1° away from the central fixation along the horizontal meridian. The participants were asked to press one of two buttons to indicate whether the two characters were the same or not while remained fixation on the center fixation point throughout the trial. Feedback was provided through an auditory cue for both correct and incorrect trials. The matching performance for each character type was measured at seven luminance contrast levels. There were 40 trials in each block for each contrast level. Each block contains one character type. The experiment was composed of 10 blocks. Each participant was presented with 2800 trials in the psychophysical experiment (10 types of stimuli x 7 contrast thresholds x 40 trials). The presentation order within a block was randomized and the order between blocks was counterbalanced across participants.

Functional Magnetic Resonance Image (fMRI)

Participants. Fourteen healthy volunteers (7 males, 7 females) between 18 and 36 year old participated in this study. All the participants were right-handed, as assessed by the Edinburgh handedness inventory (Oldfield, 1971). They were native speakers of Chinese. They were naïve to the purpose of the experiment and were compensated financially for the hours they participate in the experiment. Informed consent was

obtained from each participant before scanning began. The experiment was approved by the IRB of the National Taiwan University Hospital.

Stimuli. The stimuli were the same as those used in the psychophysical experiment, except that we used the scrambled versions of the test characters as a control to establish a baseline. The scrambled stimuli were constructed first by dividing the image of a character into a 4 by 4 grids and then randomly exchanging the position of these 16 grids.

Procedure. We used the multiple block design to measure the cortical response to the stimuli. There were eight character-type blocks (upright real, non-characters, lexical, non-lexical components, and their inverted versions) and two scrambled-character blocks in each run. Each block was consisted 20 trials and lasted 20 s. Each trial includes a stimulus presentation for 500 ms, followed by a 500 ms blank period. In each run, each block was repeated three times and the order of the blocks was randomized. Each run was presented four times for each participant. To keep the participants' attention to the stimuli, one used a one-back matching task in which the participants were instructed to press a response key when the currently presented image physically matched the previous one (Kanwisher, et al., 1996).

Localizer. For each observer, we identified regions of interest (ROIs) with localizer experiments. We separated the LO into two subregions: the visual word form area (VWFA) and the fusiform face area (FFA). VWFA contained voxels showing a differential BOLD activation to real characters and to non-characters. FFA contained voxels showing a differential BOLD activation to the images of human faces and to the phase-scrambled versions of the same images (C. C. Chen, et al., 2007). LO was

defined by the overlap between anatomical structures and functional activations of the differential BOLD activations to pictures of common objects (objects that can be seen in everyday life) and to the scrambled versions of the same images (Kourtzi & Kanwisher, 2001). The OPCA was identified with the kinetic occipital region in which the 12 s epochs of motion boundary stimuli were alternated with 12 s epochs of transparent motion stimuli (Dupont et al., 1997; Zeki, Perry, & Bartels, 2003). We used the standard procedure of rotating wedge and expanding ring with black-and-white checkerboards to identify the boundaries in early visual area (Engel, et al., 1997). Thus, the V1 was identified from the polar angle and the eccentricity measurement. A voxel was considered to have a significant activation if its *t*-test had a *p*-value smaller than the α -level of .005 (uncorrected). According to this criterion, the voxel size of each ROI was varied across each observer. Above the criterion, one observer had no activation in the left VWFA for the VWFA localizer. All subsequent ROIs analyses were performed using the independent data set from the experimental runs.

Data acquisition. All images were acquired with a 3T Bruker MRI scanner located at the Interdisciplinary MRI Laboratory at the National Taiwan University. A high-resolution anatomical (T1-weighted) MRI volume scan of the entire head was conducted once on each participant (voxel size = 1 x 1 x 1 mm). Within each scanning session, both functional (T2*-weighted, BOLD) responses and anatomical images were acquired in identical planes. The images were collected in 24 transverse planes parallel to the AC-PC (anterior commissure – posterior commissure) line to cover the whole visual cortex. An echo-planar imaging sequence (Stehling, et al., 1991) was used to acquire the functional data (TR = 2000 ms, TE = 33 ms, flip angle = 90°, voxel resolution = 3.75 x 3.75 x 3 mm). The main experiment lasted 606 s (303 images). The

first 9 s (3 images) were excluded from further analyses to avoid the start-up transient. Thus, the data analyzed for each scan spanned 600 s (300 images).

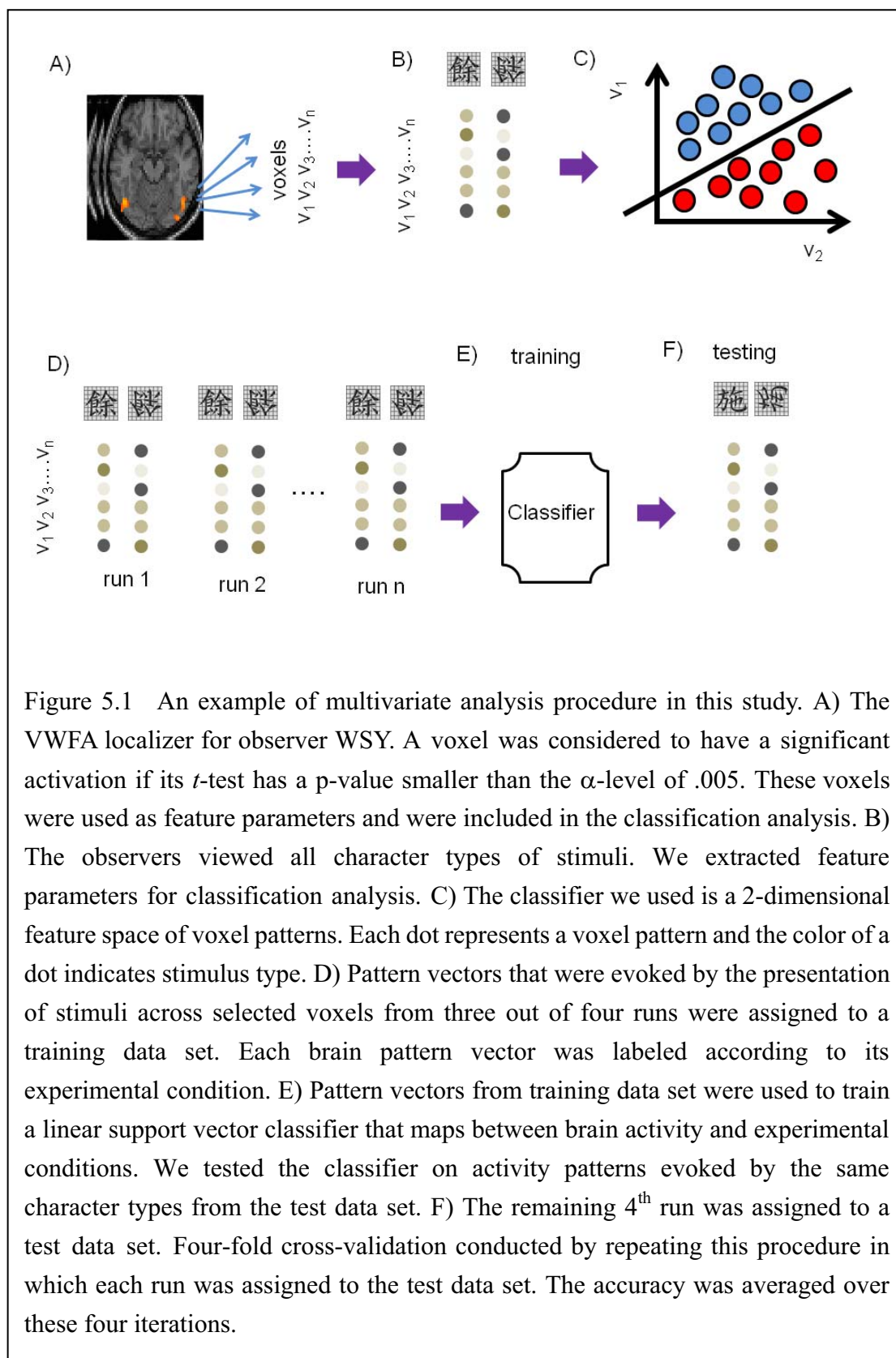
All stimuli were delivered with a MR-compatible goggle system (Resonance Technology, USA) mounted on the head of the observers. The resolution of the goggle system is 800x600 with a dot size of .096° visual angle. The frame rate is 60Hz. All the stimuli were generated on a PC-compatible computer with the Psychophysics toolbox (Brainard, 1997) under the MATLAB (The Mathworks, Matick, MA, USA) environment.

fMRI preprocessing. For each observer, we used SPM (Wellcome Department of Imaging Neuroscience, University College London) to correct head-motion artifacts and slice scan time. Then we removed the linear trend of each functional run. No spatial smoothing was used on the functional runs. Functional runs were aligned to each observer's corresponding anatomical scan and normalized into MNI space. All analyses were conducted independently for each ROI and performed with individual observer.

Multivoxel pattern classification analysis. An example of the basic framework for the classification analysis was demonstrated in Figure 5.2. For each observer, we first identified ROIs from the result of the localizers. For instance, Figure 5.2A shows the VWFA of one observer. The activation of each voxel was normalized to the z-score by the mean and the standard deviation of that voxel within a run. This normalization minimized the baseline activation difference across runs (Figure 5.2B). The normalized activation then fed to a support vector machine algorithm (LIBSVM, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>), with voxels as “features”, to determine the best parameters for classification (Figure 5.2C). Notice that, to account for the slow

hemodynamic responses, the first time point of each block was not included in the analysis. We used leave-one-out method to validate of the classification parameters (Kohavi, 1995). That is, three of runs were treated as training patterns and the fourth one as a testing pattern. Totally, there were four cross-validations for each observer and each classification (Figure 5.2D to F). The accuracy was averaged over these four cross-validations. The mean prediction accuracy across all participants on each ROI was then tested against the chance level (0.5 for a two-way classification) with a one-sample *t*-test.





5.2 Results

Psychophysics

Figure 5.3 shows the proportion of correct responses as a function of contrast. For all conditions, the proportion of correct responses increased with contrast until they reached an asymptotic level. Such data was fit into the function

$$p(x) = \gamma + (1-\gamma) \times \rho \times \Phi(x, \mu, \sigma) \quad (1)$$

where $p(x)$ was the proportion of correct responses at contrast level x ; γ was the guessing factor for 2AFC and was fixed at 0.5; ρ determined the asymptotic level of the psychometric function, $\Phi(\mu, \sigma)$ is the Gaussian cumulative distribution function with the location parameter (“mean”) μ and the scale parameter (“standard deviation”) σ . The smooth curves in Figure 5.3 were fits of this function. The root of the mean squared error (RMSE) of the model fit was about .040, which is close to the mean of the standard error of measurement, .029. The values of the parameters are shown in Table 5.1.

Table 5.1 The resulting model fits to combined data set of fifteen observers. The parameter m is the mean of the contrast, σ and r is the scale parameter, and RMSE indicates the goodness-of-fit.

	μ	σ	ρ	RMSE
Upright real character	-19.895	3.42	41.77	0.0132
Inverted real character	-17.634	5.21	32.20	0.0166
Upright non-character	-17.675	3.56	32.13	0.0136
Inverted non-character	-18.162	5.50	28.78	0.0109
Upright lexical component	-19.412	3.95	42.56	0.014
Inverted lexical component	-17.284	7.03	38.11	0.0063
Upright non-lexical component	-19.249	2.71	41.46	0.0107
Inverted non-lexical component	-18.753	4.67	36.70	0.0281
Upright <i>Jiagu</i> character	-21.882	3.20	29.08	0.03
Inverted <i>Jiagu</i> character	-20.429	4.06	31.09	0.0048

At high contrasts, the proportion of correct responses for matching two upright real characters was significantly better than that for matching two inverted real characters at .05 α -level (denoted by symbol ‘*’ in Figure 5.3 A). The psychometric function for upright real characters was asymptotic to the percentage correct level of 91%, while the function for inverted characters was 81%. That is, at high contrasts, turning the real characters upside down reduced significantly the percentage correct level by about 10% ($t(10) = 3.91, p < .01$). Figures 5.3 B shows the psychometric functions for matching non-characters. There was no difference between the proportional correct responses for matching upright and matching inverted

non-characters. The psychometric function was asymptotic to the level 82% and 79% for upright and inverted non-characters. The 3% drop in performance was not statistical significance ($t(10) = 1.36, p > .05$). Thus, there is no inversion effect for non-characters. Figure 5.3 C shows the psychometric functions for matching lexical components. The psychometric function for upright lexical components was asymptotic to the percentage correct level of 92%, while the function for inverted lexical components was 86%. The matching performance for the upright components was significantly better than that for matching two inverted lexical components ($t(10) = 2.54, p < .05$). Figure 5.3 D show the psychometric functions for matching non-lexical components. The psychometric function for upright non-lexical components was asymptotic to the percentage correct level of 91%, while the function for inverted non-lexical components was 87%. That is, at high contrasts, the proportional correct responses for matching two upright non-lexical components were significantly better than those for matching two inverted non-lexical components ($t(10) = 2.64, p < .05$). Figure 5.3 E shows the psychometric functions for matching *Jiagu* characters. The psychometric function was asymptotic to the level 80% and 81% for upright and inverted *Jiagu* characters. There was no difference between the proportional correct responses for matching upright and matching inverted *Jiagu* characters ($t(10) = -0.83, p > .05$). Thus, there is no inversion effect for *Jiagu* characters. The results show that the characters and their component contain spatial configuration processing.

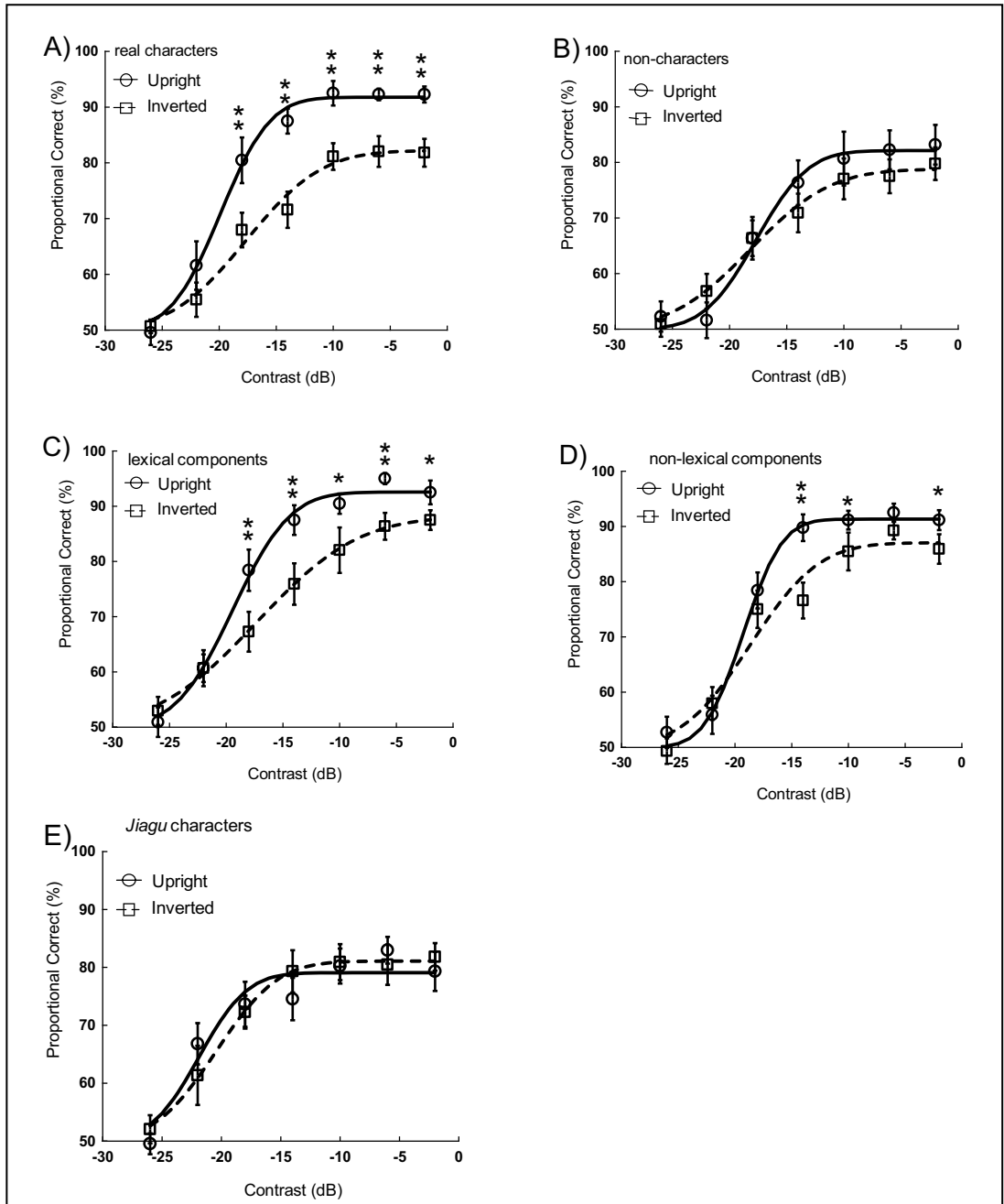


Figure 5.2 The Gaussian accumulative distribution functions as a fitting model for the data. Proportional correct responses of upright and inverted real, non-, lexical, non-lexical, and *Jiagu* characters at different contrast thresholds for 11 observers. a) real characters, b) non-characters, c) lexical components, d) non-lexical components, e) *Jiagu* characters. In each panel, the solid denotes the proportion of correct responses across all observers' data. * $p < .05$ and ** $p < .01$. Error bars are the SEM across all observers.

The multivoxel classification performance for character types

For each observer, we identified five ROIs in each hemisphere: V1, VWFA, FFA, LO, and OPCA (see Figure 5.4). Table 5.2 lists the Talairach coordinates of ROIs of the activated voxels.

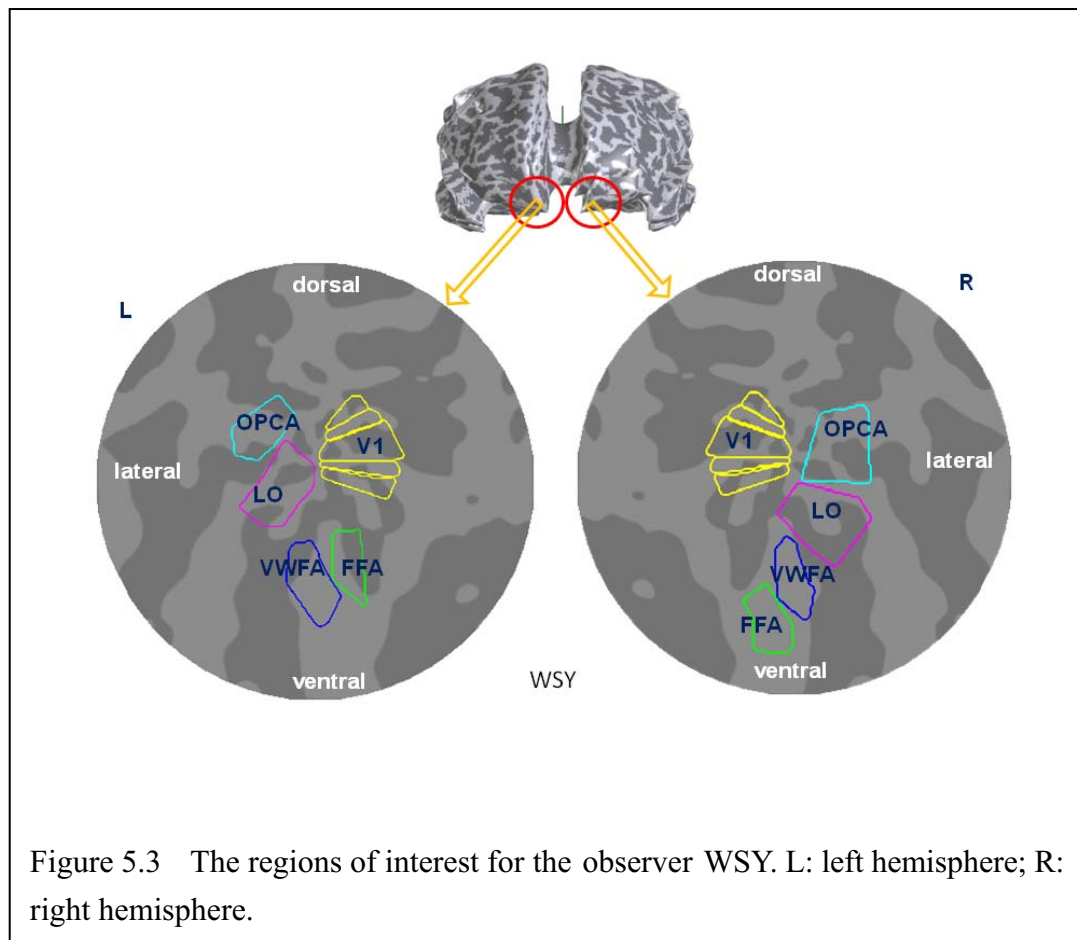


Table 5.2 The Talairach Coordinate of ROIs in this study.

Brain Regions	Hemisphere	Talairach Coordinate		
		X	Y	Z
VWFA	left	-30	-48	-17
	right	45	-69	-12
FFA	left	-33	-48	-18
	right	36	-54	-21
LO	left	-36	-72	-18
	right	42	-72	-9
OPCA	left	-45	-72	9
	right	48	-75	12

We employed a linear kernel support vector machine to train the activity of the population of voxels in each ROI with all character stimuli. We first tested whether the brain regions could distinguish the character types from scrambled lines. Figure 5.5 shows the prediction accuracies for distinguishing between characters types and scrambled lines in V1, VWFA, FFA, LO, and OPCA. The accuracy rate of 50% indicates chance performance (two-way classification) and a rate of 100% indicates that the classifier could perfectly distinguish the patterns of different conditions based on BOLD activity. We obtained significant higher prediction accuracies for all character types than chance performance in VWFA, LO, and OPCA (Figure 5.5). The results suggest that there are mechanisms responsive to visual word form information in those areas. The activation pattern of V1 was the same for character types and scrambled characters. This is consistent with the notion that V1 neurons are only selective to local contrast and line segments (Hubel & Wiesel, 1968; Livingstone,

Freeman, & Hubel, 1996). The FFA activation pattern also showed no difference between character types and scrambled characters. This result was consistent with Chapter 4 that FFA is specific for face information but not visual word form information. These results ensure that the multivoxel patterns of activity reliably represented the properties of different brain regions.



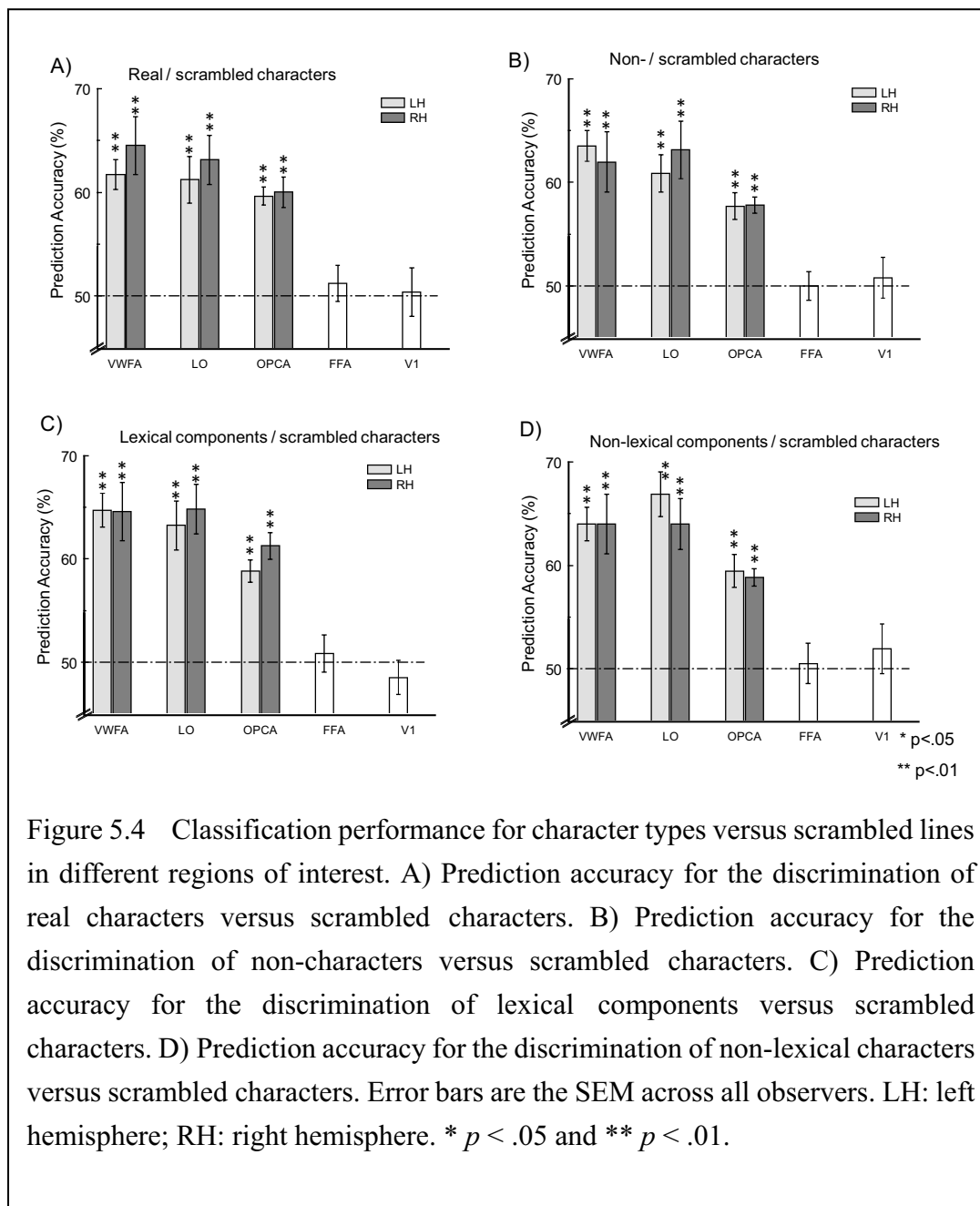
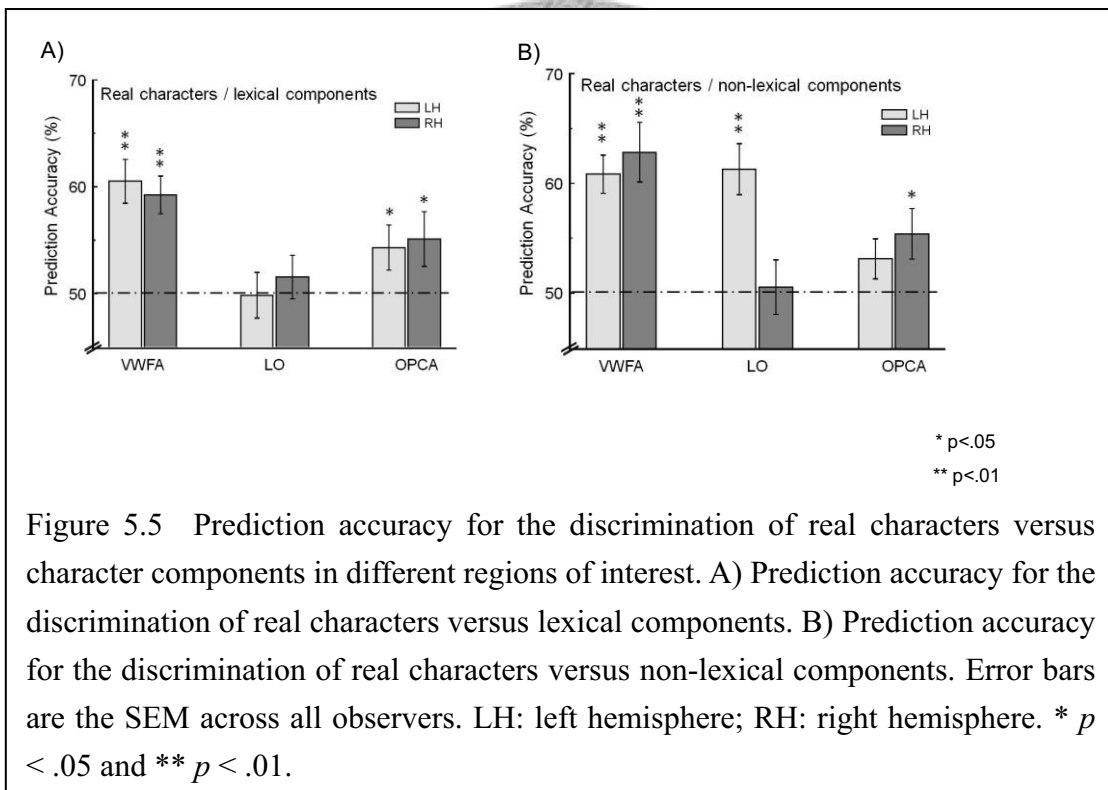


Figure 5.4 Classification performance for character types versus scrambled lines in different regions of interest. A) Prediction accuracy for the discrimination of real characters versus scrambled characters. B) Prediction accuracy for the discrimination of non-characters versus scrambled characters. C) Prediction accuracy for the discrimination of lexical components versus scrambled characters. D) Prediction accuracy for the discrimination of non-lexical characters versus scrambled characters. Error bars are the SEM across all observers. LH: left hemisphere; RH: right hemisphere. * $p < .05$ and ** $p < .01$.

Figure 5.6A shows the accuracy for classifying real characters and lexical components in VWFA (60.5% for the left hemisphere and 59.2% for the right hemisphere), LO (49.9% for the left hemisphere and 51.5% for the right hemisphere), and OPCA (54.3% for the left hemisphere and 55.1% for RH). The bilateral VWFA ($t(12) = 5.14, p = .000 < .01$ for the left hemisphere and $t(13) = 5.22, p = .000 < .01$ for the right hemisphere) is significantly higher than the chance level. Figure 5.6B shows

the accuracy for classifying real characters and non-lexical components in VWFA (60.8% for the left and 62.9% for the right hemisphere), LO (61.3% for the left and 50.6% for the right hemisphere), and OPCA (53.1% for the left and 55.4% for the right hemisphere). The bilateral VWFA ($t(12) = 6.24, p = .000 < .01$ for the left and $t(13) = 4.72, p = .000 < .01$ for the right hemisphere), the left LO ($t(13) = 4.88, p = .001 < .01$), and the right OPCA ($t(13) = 2.35, p = .02 < .05$) show significant higher accuracy for classifying real characters and non-lexical components. The results suggest the bilateral VWFA could distinguish a character and its components.



The classification performance correlated with inversion effects

Figure 5.7A shows the accuracy for classifying upright and inverted real characters in VWFA (60.9% for the left hemisphere and 53.9% for the right hemisphere), LO (57.2% for the left hemisphere and 50.4% for the right hemisphere) and OPCA (55.4% for the left hemisphere and 50.7% for the right hemisphere). The bilateral VWFA ($t(12) = 7.9, p = .000 < .01$ for the left and $t(13) = 2.16, p = .03 < .05$ for the right hemisphere), the left LO ($t(13) = 4.97, p = .001 < .01$), and the left OPCA ($t(13) = 4.49, p = .000 < .01$) show significantly higher accuracy than the chance performance. This implies that the left occipital regions are sensitive to the spatial configuration of a character. Figure 5.7B shows the accuracy for classifying upright and inverted non-characters in VWFA (50% for the left and 52% for the right hemisphere), LO (50% for the left and 51.3% for the right hemisphere) and OPCA (48.2% for the left and 51.5% for the right hemisphere). There was no significant accuracy from the chance performance in all regions ($p > .05$). This result showed no differential activation between upright and inverted non-characters in the occipital regions. Figure 5.7C shows the accuracy for classifying upright and inverted lexical components in VWFA (56.5% for the left and 55.3% for the right hemisphere), LO (53.9% for the left and 56% for the right hemisphere), and OP (49.9% for the left and 52.2% for the right hemisphere). The VWFA ($t(12) = 5.53, p = .0001 < .01$ for the left and $t(13) = 2.88, p = .000 < .01$ for the right hemisphere), and the right LO ($t(13) = 4.3, p = .000 < .01$) show a significantly higher accuracy for classifying upright and inverted components. Figure 5.7D shows the accuracy for classifying upright and inverted non-lexical components in the VWFA (50.5% for the left and 58.3% for the right hemisphere), LO (50.1% for the left and 58.8% for the right hemisphere), and OP (50.7% for the left and

50.3% for the right hemisphere). The right VWFA ($t(13) = 4.47, p = .0003 < .01$) and right LO ($t(13) = 5.66, p = .000 < .01$) show a significantly higher accuracy for classifying upright and inverted components. This result suggests the right ventral occipital region may be responsible for the local configuration of a character.

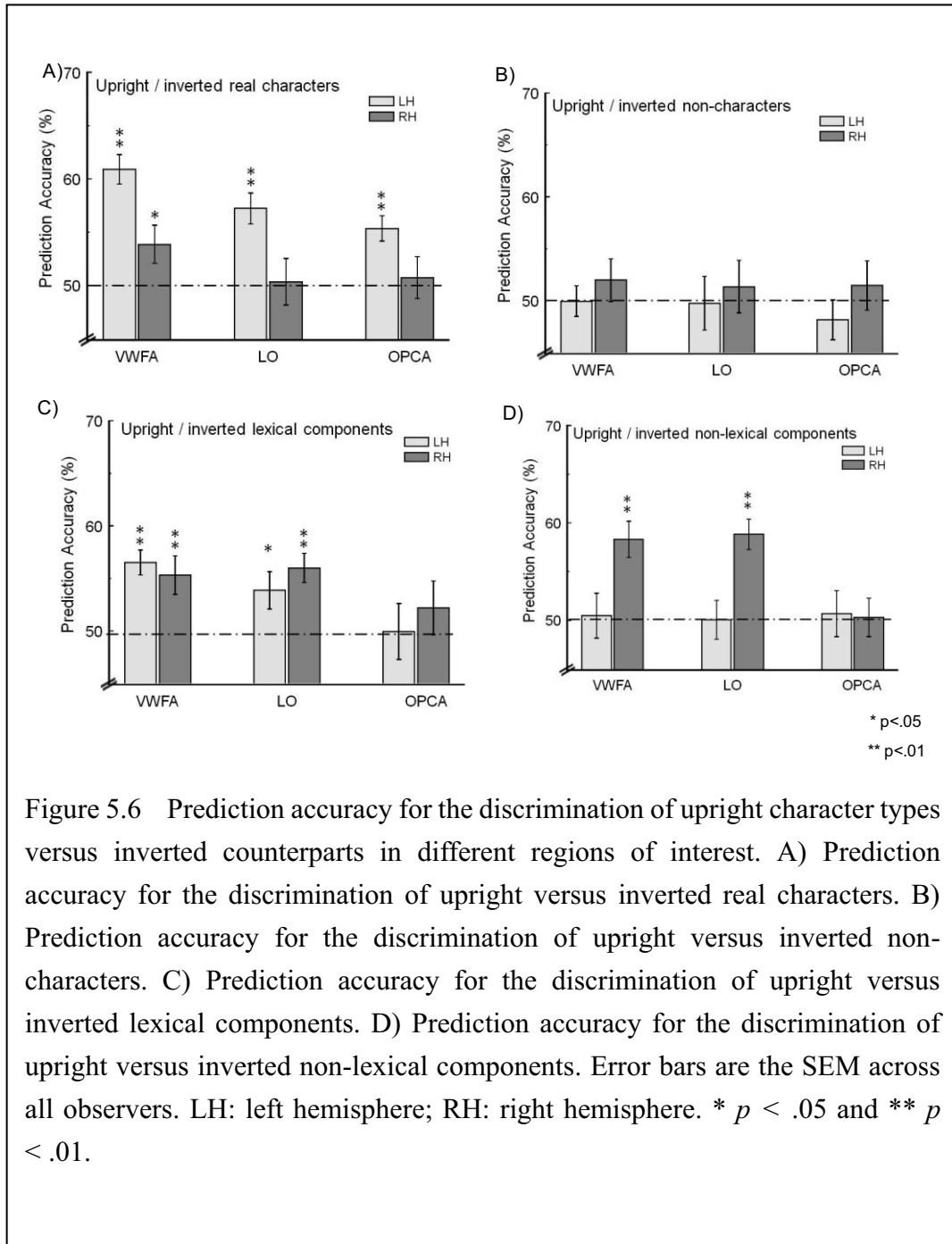


Figure 5.6 Prediction accuracy for the discrimination of upright character types versus inverted counterparts in different regions of interest. A) Prediction accuracy for the discrimination of upright versus inverted real characters. B) Prediction accuracy for the discrimination of upright versus inverted non-characters. C) Prediction accuracy for the discrimination of upright versus inverted lexical components. D) Prediction accuracy for the discrimination of upright versus inverted non-lexical components. Error bars are the SEM across all observers. LH: left hemisphere; RH: right hemisphere. * $p < .05$ and ** $p < .01$.

5.3 Discussion

Inversion effect and spatial configuration

In the psychophysical experiment, we found that the percentage correct by observers matching the upright real characters was greater than that for their inverted versions. In addition, the percentage correct in matching the upright non-characters was similar to that for their inverted versions. This inversion effect is a signature of the spatial configuration processing in the visual analysis of well-practiced objects (Carey & Diamond, 1977; Leder & Bruce, 2000; Rhodes, et al., 1993; Yin, 1969). Such inversion effect is also observed in the lexical and the non-lexical components. The results show that there are spatial configuration processings in both the local and the global stages. That is, the visual system analyzes not only the spatial relationship among strokes in a character component but also the spatial relationship among components in a character. Furthermore, while the *Jiagu* characters have the same spatial configuration as modern Chinese characters, no inversion effect was observed in them. This absent of inversion effect may result from a lack of familiar components in a *Jiagu* character

The spatial configuration processing of a visual word may differ from that of a face. While previous studies found that faces show the inversion effect, the local features of a face, such as eyes, mouths, and noses, did not show the significant inversion effect (Riesenhuber & Wolff, 2009). Such inversion effect implies that spatial configuration only occurs in the “whole face”. Our results show that spatial configuration occurs in both local and global levels of visual word form processing.

This difference implies that the mechanism for characters is different from the mechanism for face processing. Taken together, our result supported the hierarchical hypothesis of visual word form processing, in which a word was recognized through an analysis of the local configuration and the global configuration, rather than an analysis of its components.

In this study, we obtained a robust inversion effect for Chinese characters. This effect was strong in real characters but was absent in the non-characters. The results were consistent with familiar object recognition literature, showing that expert processing is tuned by extensive practice (Gauthier & Tarr, 1997). The acquisition of expertise is a process which involves extensive practice over a period of years. It is possible that if the co-occurrence for individual features at specific positions occurs with great frequency, the integration among these features could become stronger at the positions (Hebb, 1949). If features do not often co-occur, the integration among features could become weaker or disconnected. Thus, the statistical learning mechanism may be an explanation for the sensitivity of spatial configuration that we found in orthographic stimuli. In addition, the spatial configuration was also involved in the character components that we found the inversion effect for lexical and non-lexical components. It is possible that the integration of strokes of a component may be due to the simple exposure of the same components recurring in different characters. Chen, Allport, and Marshall (1996) demonstrated that for a fluent Chinese reader, with increases in the number of constituents in a character, the matching performance for two characters becomes worse. Furthermore, they also reported that non-Chinese readers showed a similar pattern as Chinese readers, in matching two character conditions after a few hundred exposures to composite Chinese characters.

These results indicate that visual analysis of a composite character is also reliant on the orthographic constituents (Y. P. Chen, et al., 1996; Tsang & Chen, 2009).

Functional distinctions between the left and right hemispheres in visual word form processing

We used the multivoxel pattern analysis to characterize the classification performance for different types of characters of various regions of interest. The extrastriate cortex was able to distinguish between character types and scrambled characters, while the early visual area (V1) failed to classify character types. That is, the extrastriate cortex contains character information that is sufficient to distinguish all character types from segments. The results confirm that visual word form processing involves a widespread activation in the extrastriate cortex, including the bilateral occipitotemporal regions, as well as the occipitotemporal and occipitoparietal regions (Chan et al., 2009; Kuo et al., 2004; Tan et al., 2000). Such analysis of a visual word only occurs in the VWFA, not in the FFA. There was no difference in the accuracy between characters and scrambled segments. That is, the function of the left occipitotemporal regions showed a certain degree of categorical specification.

The left VWFA activation was able to distinguish between the upright and inverted real characters and lexical components, but failed to discriminate non-characters or components between their inverted versions. Because real characters and lexical components can be used independently, this result suggests that VWFA is selective for global configuration in a character. In contrast to the left VWFA, the right VWFA and LO showed a high accuracy for classifying upright and inverted

components. These regions failed to classify upright characters and inverted counterparts, although the right FG showed a slight ability to discriminate between real characters and inverted counterparts. This result implies that the right extrastriate region may encode the local configuration of a character. Our findings highlight the distinctive functions of hemispheric asymmetry for visual word form processing and demonstrated for the first time that the functional segregation between the left and right hemisphere in visual word form processing. The hemispheric asymmetric was also reported in Navon figure identification (Fink, Marshall, Halligan, & Dolan, 1999; Lux, Marshall, Thimm, & Fink, 2008). They found that the left hemisphere was activated when participants identified the global level of the Navon figure and the right hemisphere was activated when participants identified the local level of the figure. However, notice that those Navon figure experiments tagged more about visual attention than image processing.

Split-field tachistoscopic studies found that the left visual field/the right hemisphere advantage is involved in Chinese orthographic processing (e.g., Tzeng, et al., 1979; Yang & Cheng, 1999). Neuroimaging studies of Chinese characters processing have shown a consistent activation in the right occipital regions for fluent readers (Tan, Feng, et al., 2001; Tan, et al., 2000). It has been found that reading print-like fonts shows the right visual field/the left hemisphere advantage and cursive fonts shows the left visual field/the right hemisphere advantage (e.g., Hellige & Adamson, 2007). Because analyses of handwriting styles may demand a greater local shape of a font than the analysis of block fonts, it has been suggested that the right hemisphere may subserve to relevant components of visual word forms and the left hemisphere involves the global shape of a word (Bryden & Allard, 1976). Further

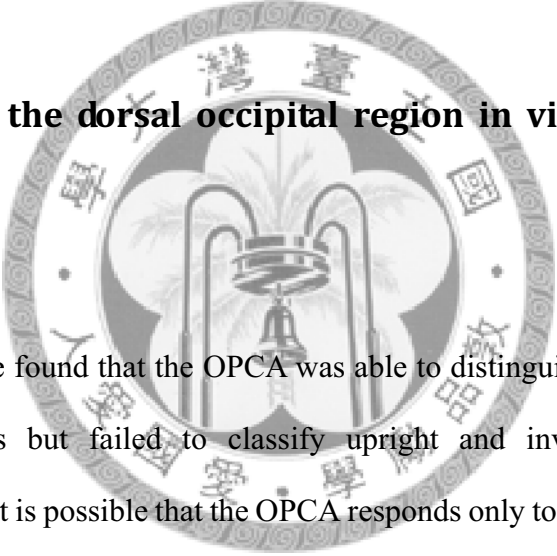
support for hemispheric asymmetry in visual word processing can be found in the left hemisphere lesion studies. Cohen and colleagues (2004) found that while children had the damage in the left occipitotemporal region, they can develop normal reading ability. Such recovery of reading ability may be due to the visual word information processing in the right hemisphere. Furthermore, several findings have shown that with the increase of reading skills, the activation of the right FG (Schlaggar et al., 2002; Turkeltaub, Gareau, Flowers, Zeffiro, & Eden, 2003) decreased. The findings in our study suggest that the efficient processing of visual word forms may rely on the integration of information between two hemispheres, in which the right hemisphere is selective to the local shape of a word and the left hemisphere is responsive for the global configuration processing (see review by Dien, 2008, 2009).

Visual word form area vs. fusiform face area

While faces and Chinese characters have distinctive physical features, researchers found that faces and characters have inversion effects, suggesting that spatial configuration processing is involved in these two types of objects (Kao et al., 2010, see Chapter 4). Numerous neuroimaging and neuropsychology studies have identified two distinct areas in the occipitotemporal region that are selective for visual word (L. Cohen, et al., 2000; L. Cohen, et al., 2002) and face information (Kanwisher, et al., 1996) and called VWFA and FFA, respectively. Both of these areas are located close to each other, in the FG. While both visual word form and face information activated bilateral FG, the lateralization was found in these two types of categories, in which the right hemisphere is strongly activated by faces (Kanwisher, et al., 1996; Rossion et al., 2000) and the left hemisphere is strongly activated by visual word forms (L. Cohen, Jobert, Le Bihan, &

Dehaene, 2004). Therefore, Rossion and colleagues (2000) argued that the right FFA is involved in holistic processing, whereas the left FFA is involved in feature analysis processing. The right hemisphere is specific to configural coding of faces that utilizes spatial relationship in an object, whereas the left hemisphere focuses on local feature analysis of faces (Carey & Diamond, 1977). The separate functional asymmetry for faces and characters may imply a hierarchical level of object perception in which the left hemisphere is dominant for orthographic processing, whereas the right hemisphere is dominant for face processing, even though contralateral hemisphere is involved in object perception (Dien, 2008, 2009).

The function of the dorsal occipital region in visual word form processing



In this study, we found that the OPCA was able to distinguish the characters and scrambled characters but failed to classify upright and inverted characters and components. Hence, it is possible that the OPCA responds only to local features such as corners, T-junctions, and L-junctions that were changed when we scrambled the characters. Furthermore, this region also showed a slightly higher accuracy for the discrimination between real characters and character components. OPCA may be sensitive to the number of strokes in a character which contains more strokes than a component. This result was consistent with previous studies which suggested the occipitoparietal region was involved in the processing of the spatial arrangement of strokes or character components (Fu, et al., 2002; Kuo, et al., 2001; Tan, Feng, et al., 2001; Tan, Liu, et al., 2001). In particular, it has been found that the right dorsal visual

stream is involved in visual word recognition. For example, a magnetoencephalography (MEG) study demonstrated that the activation of the right posterior parietal cortex occurs in 100 ms post-stimulus onset in visual word recognition tasks (Pammer, Hansen, Holliday, & Cornelissen, 2006). Neurophysiological evidence also found that the right dorsal visual stream deficits could induce difficulty in binding features of a word (Kevan & Pammer, 2009; Pammer & Vidyasagar, 2005). Convergent evidence suggests that the right OPCA is involved in early visual feature analysis of a word.

The current behavioral and fMRI results showed the visual system may analyze characters in a hierarchical way, which was consistent with Chapter 4 (Kao et al., 2010). The occipitoparietal regions may analyze the local features of a visual word form. Then, the right fusiform and lateral occipital regions may play an intermediate role in integrating the local information in a word. Finally, the left plays a critical role in analyzing global configurations in a character.

5.4 Summary

In summary, we used psychophysics and fMRI with multivoxel pattern analysis to investigate the configural processing of orthographic stimuli. The behavioral results showed the inversion effect of real characters, lexical components, and non-lexical components. This result suggested that the spatial configuration occurs in global and local configuration levels of a visual word. The multivoxel pattern analysis is successful distinguishing the representations of visual word forms in the visual cortex. While the ventral visual cortex contained character information, distinctive cortical regions play separate roles in processing characters. The left VWFA showed the discrimination between upright and inverted real characters, and upright and inverted

lexical components, while the right VWFA and LO showed the discrimination between upright and inverted lexical components and upright and inverted non-lexical components. This result suggests that the right occipitotemporal regions analyze the local configuration, while the left VWFA analyzes the global configuration. The right occipitoparietal regions showed the discrimination between real characters and components. This result implies that this region analyzes local features of a character. It is the first report that demonstrates the functional segregation between the left and right hemisphere for visual word form processing. Taken together, this study supports a hierarchical processing of visual word forms in the visual cortex.

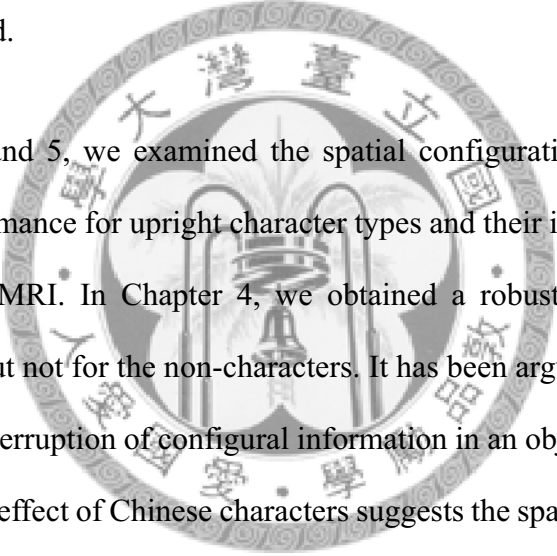


Chapter 6 : General Discussions

This study explored the processing of visual word perception and its neural network at the both behavioral and neural levels. We started with a spatial summation experiment to determine the mechanisms for character perception, using 2AFC psychophysics (Chapter 2). We investigated the performance of observers in detecting and discriminating various types of character-like stimuli as a function of visual word form size and retinal eccentricity. The detection thresholds for the same stimulus size and eccentricity were the same for all types of stimuli. This result suggests that detection is mediated by mechanisms that tune to local features but with no specialization for Chinese characters. The contrast thresholds for discriminating between real and pseudo- or non-characters were higher than for discriminating between real characters and *Jiagu* characters or scrambled characters. The result shows that the discriminability, however, requires a specific mechanism for perceiving characters. The difference between pseudo-, non-characters and *Jiagu* characters is that formers are constructed by familiar components while the later contains no familiar components. The specific mechanism for perceiving characters may rely on the global shape of a character or components information.

In Chapter 3, we further used the retinotopic paradigm with fMRI to define the visual cortical regions which produce unique representations for Chinese characters. The brain activation for characters was extended from the primary visual area to the higher visual regions in the occipital cortex. Compared with simple lines (e.g. scrambled characters), the character stimuli showed greater activation in the ventral occipital region, especially the FG, suggesting that this region is sensitive to visual

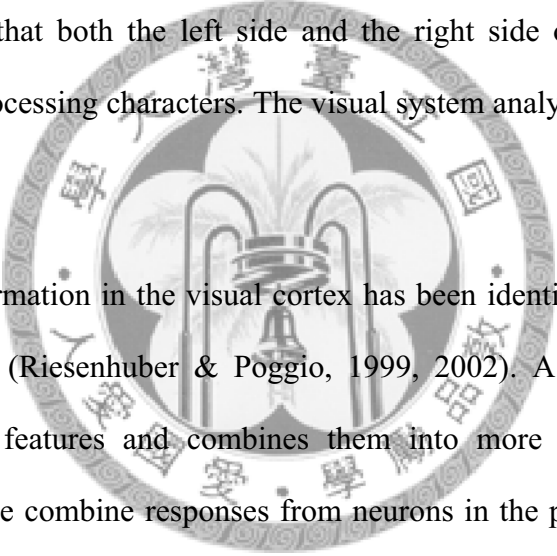
words, rather than local features. In particular, the retinotopic representation for character stimuli was in the ventral occipital region where we found that the FG is selective for the contralateral visual fields and centre-to-periphery map. This implies that the FG is involved in the spatial location information of stimuli. Furthermore, contrasted with scrambled characters, the FG showed a preference for characters presented in the fovea. Such foveal selectivity was located in the VWFA, suggesting that the VWFA is associated with the processing of central information. The retinotopic properties in the visual cortex imply that the ventral occipital region is a stimulus-driven region and the activation of this region can be attributed to the visual information of a word.



In Chapters 4 and 5, we examined the spatial configuration of a character by comparing the performance for upright character types and their inverted versions with psychophysics and fMRI. In Chapter 4, we obtained a robust inversion effect for Chinese characters but not for the non-characters. It has been argued that the inversion effect is due to the interruption of configural information in an object (Gauthier & Tarr, 1997). The inversion effect of Chinese characters suggests the spatial configuration in a character processing. Neuroimaging evidence showed that the left FG and a small area in the bilateral lateral occipital regions may be responsible for the spatial configuration processing of a character. In Chapter 5, we also found that the spatial configuration was also involved in the character components that we found the inversion effect for lexical and non-lexical components. The integration of strokes in a component may be due to the simple exposure of the same components recurring in different characters.

In order to reveal the different brain regions for processing different spatial configurations in a character, we used fMRI with multivariate methods (e.g. SVM) to

classify the representations of character and component information in the visual cortex. The left VWFA showed the discrimination between upright and inverted real characters, and upright and inverted lexical components, while the right occipitotemporal regions showed the discrimination between upright and inverted lexical components and upright and inverted non-lexical components. This result suggests that the left VWFA is selective for spatial relations among components in a character, while the right occipitotemporal regions encode spatial relationship among strokes. The right occipitoparietal regions showed the discrimination between real characters and components. This result implies that this region analyzes local features of a character. The results suggest that both the left side and the right side of visual cortex play important roles in processing characters. The visual system analyzes visual word form in a hierarchical way.



The visual information in the visual cortex has been identified as a hierarchical processing from V1 (Riesenhuber & Poggio, 1999, 2002). A hierarchical process begins with simple features and combines them into more complicated objects. Neurons in each stage combine responses from neurons in the preceding stages. Our behavioral evidence revealed that the mechanism of visual word perception is based on the feature encoding and then, in turn, of the analysis of the spatial configuration in a word. Neuroimaging evidence demonstrated that a hierarchical processing of visual word forms in the visual cortex. The visual system may analyze a visual word in a hierarchical way. Such hierarchical processing was also found in studies with other visual stimuli. Oswald et al. (2008) showed that the activations in the early visual areas were influenced by the local features, such as dipole orientation, of a Glass pattern (Glass, 1969), while the lateral occipitotemporal regions were engaged by the global

forms of Glass patterns (e.g., Kourtzi & Huberle, 2005). Chen et al. (2007) showed that the activations in the FFA were influenced by the global configuration of a face, or the symmetry, while the activations in the dorsal occipital cortex were only influenced by local features. Furthermore, Cohen et al. (2003) proposed a model of visual word form perception in the visual system in which the early visual areas analyzed letters in contralateral visual fields and then bilateral FG computed the visual word form representations. These studies drew similar conclusions to ours with regard to visual word form. We can conclude that the flow of the ventral stream, from local to global, flows from posterior to anterior and from dorsal to ventral on the visual cortical surfaces.

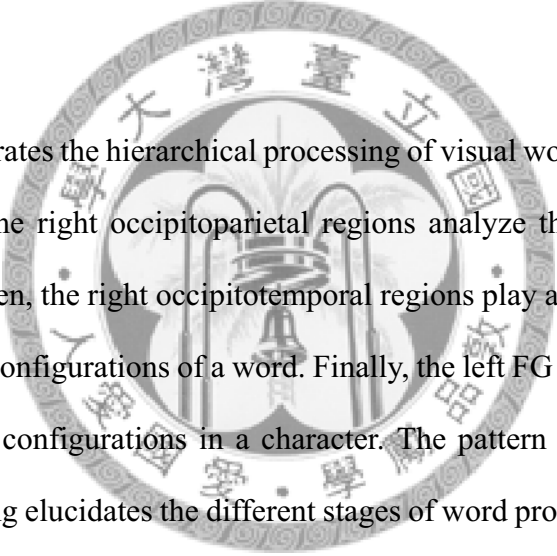
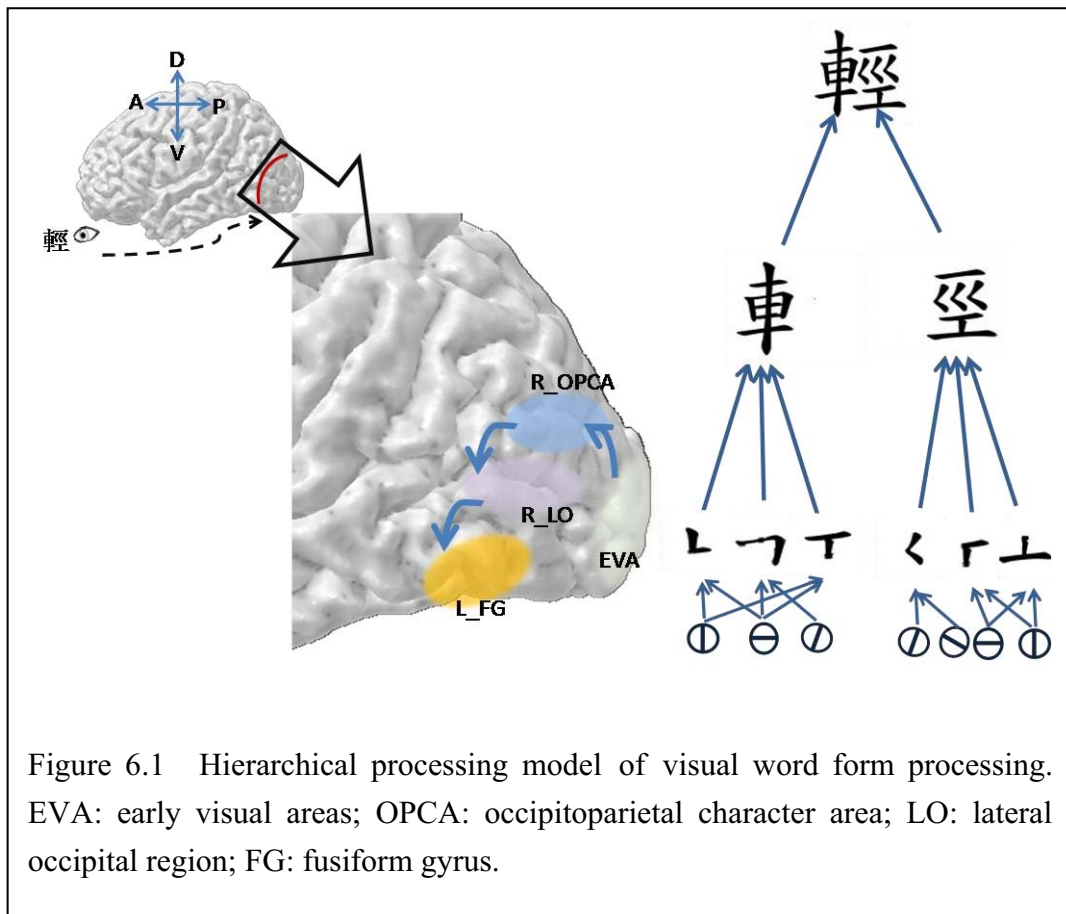


Figure 6.1 illustrates the hierarchical processing of visual word form processing in the visual system. The right occipitoparietal regions analyze the local features of a visual word form. Then, the right occipitotemporal regions play an intermediate role in integrating the local configurations of a word. Finally, the left FG plays a critical role in analyzing the global configurations in a character. The pattern of neural activity for visual word processing elucidates the different stages of word processing are integrated into visual processing and influence human behavior. Based on these findings, this study makes a significant contribution toward our understanding of visual word form processing in the visual system and provides a solid foundation for future studies.



Future directions

We have addressed how our visual system analyzes visual word forms and suggested a hierarchical processing for visual word form perception. There are some related issues that need to be solved. The examination of these issues would enable us to build a more complete framework for understanding the processing of visual word perception.

The feedback processing of hierarchical model of visual word forms

In the current study, we clearly demonstrated, from dorsal to ventral, that the visual system analyzes objects in a hierarchical way. The hierarchical model could account the results of visual processing of Chinese characters. Recently, researchers

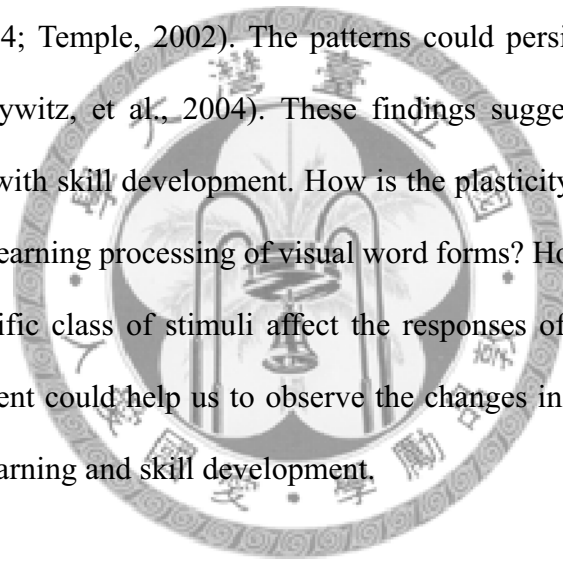
argued that a feedforward pathway may provide some useful information in explaining the complexity of visual perception. The feedback connections play a role in integrating visual information from different cortical regions to generate an integrate perception (Bullier, 2001). However, the relative contributions of feedforward and feedback connectivities are still controversial. Future works are needed to explore the role of feedback in object processing.

The temporal hierarchy of visual word form processing

In this thesis, we highlighted the distinctive roles of visual regions in processing visual word forms and supported a hierarchical processing of visual word forms. However, due to the limitation of temporal resolutions in fMRI, we are unsure of the time sequence of visual word form processing across different visual cortices. Previous event-related potential (ERP) studies demonstrated that, under object recognition processing, the neural activity of local features in an object occurs about 135 ms and then the object recognition occurs about 150-300 ms after image onset (Johnson & Olshausen, 2003). Furthermore, Schendan and Lucia (2010) used ERP to identify the temporal dynamic of different object sensitive cortical regions. They found that, compared with phase-scrambled objects, when participants recognize a common object, the processing of objects involved different time courses of the figure-ground segregation (95-175 ms), perceptual grouping (200-300 ms), and decision making (200-530 ms). The results showed the early peak activity located in the early visual areas and later peak activity located in the higher visual regions. Understanding the time courses of different cortical regions of visual word form processing could elucidate how the local and global configurations in a word interact with each other.

The development of word perception skills

While there have been many neurobiological studies of the development of reading (B. A. Shaywitz et al., 2002; S. E. Shaywitz et al., 2003; Turkeltaub, et al., 2003), there is no generally accepted theory of how the cortex develops with skill yet. In this thesis, we found that the functions of ventral occipital regions may be modulated by visual perceptual experience. For example, longitudinal studies of reading ability found that, with the increases of reading skills, activation of VWFA was increased for 9-10 year old children (B. A. Shaywitz, et al., 2002; S. E. Shaywitz, et al., 2003). Similarly, neuroimaging studies found that the increases of activation in VWFA when the poor and impaired readers were with intensive training with reading skills (B. A. Shaywitz et al., 2004; Temple, 2002). The patterns could persist for 1-3 years after training (B. A. Shaywitz, et al., 2004). These findings suggest that the VWFA is strongly associated with skill development. How is the plasticity of particular regions associated with the learning processing of visual word forms? How does the frequency of exposure to specific class of stimuli affect the responses of cortical cortex? The impact of development could help us to observe the changes in the cortex associated with orthographic learning and skill development.





References

- Academia Sinica balanced corpus. (Version 3) [CDROM] (1998). from Academia Sinica:
- Arcaro, M. J., McMains, S. A., Singer, B. D., & Kastner, S. (2009). Retinotopic organization of human ventral visual cortex. *Journal of Neuroscience*, *29*(34), 10638-10652.
- Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(21), 9087-9092.
- Barlow, H. B. (1958). Temporal and spatial summation in human vision at different background intensities. *The Journal of Physiology*, *141*(2), 337-350.
- Battista, J., Kalloniatis, M., & Metha, A. (2005). Visual function: the problem with eccentricity. *Clinical and Experimental Optometry*, *88*(5), 313-321.
- Baumgardt, E. (1959). Visual spatial and temporal summation. *Nature*, *184*(Suppl 25), 1951-1952.
- Beauregard, M., Chertkow, H., Gold, D., Karama, S., Benhamou, J., Babins, L., et al. (1997). Word priming with brief multiple presentation technique: Preservation in amnesia. *Neuropsychologia*, *35*(5), 611-621.
- Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., et al. (2003). Neural correlates of lexical access during visual word recognition. *Journal of Cognitive Neuroscience*, *15*(3), 372-393.
- Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology*, *203*(1), 237-260.

- Bonda, E., Petrides, M., Frey, S., & Evans, A. (1995). Neural correlates of mental transformations of the body-in-space. *Proceedings of the National Academy of Sciences of the United States of America*, 92(24), 11180-11184.
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226(5241), 177-178.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433-436.
- Brewer, A. A., Liu, J., Wade, A. R., & Wandell, B. A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nature Neuroscience*, 8(8), 1102-1109.
- Bryden, M. P., & Allard, F. (1976). Visual hemifield differences depend on typeface. *Brain and Language*, 3(2), 191-200.
- Bullier, J. (2001). Integrated model of visual processing. *Brain Research. Brain Research Reviews*, 36(2-3), 96-107.
- Campbell, F. W., & Robson, J. G. (1964). The attenuation characteristics of the visual system determined by measurements of flicker threshold, brightness and pupillomotor effect of modulated light. *Documenta Ophthalmologica*, 18, 83-84.
- Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, 195(4275), 312-314.
- Cattell, J. (1886). The time taken up by cerebral operations. *Mind: A Quarterly Review of Philosophy*, 11, 277-282, 524-538.
- Chan, S. T., Tang, S. W., Tang, K. W., Lee, W. K., Lo, S. S., & Kwong, K. K. (2009). Hierarchical coding of characters in the ventral and dorsal visual streams of Chinese language processing. *Neuroimage*, 48(2), 423-435.
- Chen, C. C., Kao, K. L., & Tyler, C. W. (2007). Face configuration processing in the human brain: the role of symmetry. *Cerebral Cortex*, 17(6), 1423-1432.

- Chen, Y. P., Allport, D. A., & Marshall, J. C. (1996). What are the functional orthographic units in Chinese word recognition: The stroke or the stroke pattern? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *49*, 1024-1043.
- Cheng, C. M., & Yang, M. J. (1989). Lateralization in the visual perception of Chinese characters and words. *Brain and Language*, *36*(4), 669-689.
- Chung, S. T. (2002). The effect of letter spacing on reading speed in central and peripheral vision. *Investigative Ophthalmology and Visual Science*, *43*(4), 1270-1276.
- Chung, S. T., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research*, *41*(14), 1833-1850.
- Chung, S. T., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research*, *38*(19), 2949-2962.
- Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word form area. *Neuroimage*, *22*(1), 466-476.
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123* (Pt 2), 291-307.
- Cohen, L., Jobert, A., Le Bihan, D., & Dehaene, S. (2004). Distinct unimodal and multimodal regions for word processing in the left temporal cortex. *Neuroimage*, *23*(4), 1256-1270.
- Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, *125*(Pt 5), 1054-1069.
- Cohen, L., Martinaud, O., Lemer, C., Lehericy, S., Samson, Y., Obadia, M., et al. (2003). Visual word recognition in the left and right hemispheres: anatomical and functional correlates of peripheral alexias. *Cerebral Cortex*, *13*(12), 1313-1333.

- Cohen, M. S., Kosslyn, S. M., Breiter, H. C., DiGirolamo, G. J., Thompson, W. L., Anderson, A. K., et al. (1996). Changes in cortical activity during mental rotation. A mapping study using functional MRI. *Brain, 119 (Pt 1)*, 89-100.
- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning, 20*, 273-297.
- Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage, 19(2 Pt 1)*, 261-270.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9(2)*, 179-194.
- Davidoff, J., Fonteneau, E., & Fagot, J. (2008). Local and global processing: Observations from a remote culture. *Cognition, 108(3)*, 702-709.
- DeFrancis, J. (1984). *The Chinese language: Fact and fantasy*. Honolulu: University of Hawaii Press.
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences, 9(7)*, 335-341.
- Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., et al. (2004). Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. *Psychological Science, 15(5)*, 307-313.
- Dehaene, S., Le Clec, H. G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *Neuroreport, 13(3)*, 321-325.
- DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., et al. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America, 93(6)*, 2382-2386.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special - an effect of expertise. *Journal of Experimental Psychology:General, 115(2)*, 107-117.

- Dien, J. (2008). Looking both ways through time: The Janus model of lateralized cognition. *Brain and Cognition*, 67(3), 292-323.
- Dien, J. (2009). A tale of two recognition systems: implications of the fusiform face area and the visual word form area for lateralized object recognition models. *Neuropsychologia*, 47(1), 1-16.
- Dong, Y., Fukuyama, H., Honda, M., Okada, T., Hanakawa, T., Nakamura, K., et al. (2000). Essential role of the right superior parietal cortex in Japanese kana mirror reading: An fMRI study. *Brain*, 123 (Pt 4), 790-799.
- Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., & Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision*, 3(10), 586-598.
- Dupont, P., De Bruyn, B., Vandenberghe, R., Rosier, A. M., Michiels, J., Marchal, G., et al. (1997). The kinetic occipital region in human visual cortex. *Cerebral Cortex*, 7(3), 283-292.
- Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, 7(2), 181-192.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598-601.
- Esterman, M., Chiu, Y. C., Tamber-Rosenau, B. J., & Yantis, S. (2009). Decoding cognitive control in human parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 106(42), 17974-17979.
- Everson, M. E. (1998). Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *Modern Language Journal*, 82(2), 194-204.
- Farah, M. J. (1990). *Visual agnosia*. Cambridge, MA: MIT Press.
- Farah, M. J., & Wallace, M. A. (1991). Pure Alexia as a Visual Impairment - a Reconsideration. *Cognitive Neuropsychology*, 8(3-4), 313-334.

- Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3), 914-921.
- Fink, G. R., Marshall, J. C., Halligan, P. W., & Dolan, R. J. (1999). Hemispheric asymmetries in global/local processing are modulated by perceptual salience. *Neuropsychologia*, 37(1), 31-40.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2), 195-207.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2, 189-210.
- Fu, S., Chen, Y., Smith, S., Iversen, S., & Matthews, P. M. (2002). Effects of word form on brain processing of written Chinese. *Neuroimage*, 17(3), 1538-1548.
- Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *Journal of Cognitive Neuroscience*, 11(4), 349-370.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191-197.
- Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research*, 37(12), 1673-1682.
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience*, 2(6), 568-573.
- Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15(4), 870-878.

- Glass, L. (1969). Moire effect from random dots. *Nature*, 223(5206), 578-580.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20-25.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, 16(2), 152-160.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41(10-11), 1409-1422.
- Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping*, 6(4), 316-328.
- Han, S., Jiang, Y., & Gu, H. (2004). Neural substrates differentiating global/local processing of bilateral visual inputs. *Human Brain Mapping*, 22(4), 321-328.
- Han, S., Weaver, J. A., Murray, S. O., Kang, X., Yund, E. W., & Woods, D. L. (2002). Hemispheric asymmetry in global/local processing: Effects of stimulus position and spatial frequency. *Neuroimage*, 17(3), 1290-1299.
- Hanson, S. J., Matsuka, T., & Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? *Neuroimage*, 23(1), 156-166.
- Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S., & Malach, R. (2003). Face-selective activation in a congenital prosopagnosic subject. *Journal of Cognitive Neuroscience*, 15(3), 419-431.
- Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, 34(3), 479-490.

- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425-2430.
- Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews: Neuroscience*, *7*(7), 523-534.
- Hayworth, K. J., & Biederman, I. (2006). Neural evidence for intermediate representations in object recognition. *Vision Research*, *46*(23), 4024-4031.
- Healy, A. F., & Cunningham, T. F. (1992). A developmental evaluation of the role of word shape in word recognition. *Memory and Cognition*, *20*(2), 141-150.
- Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley & Sons.
- Hellige, J. B., & Adamson, M. M. (2007). Hemispheric differences in processing handwritten cursive. *Brain and Language*, *102*(3), 215-227.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*, 106-154.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215-243.
- Hubel, D. H., & Wiesel, T. N. (1969). Visual area of the lateral suprasylvian gyrus (Clare-Bishop area) of the cat. *The Journal of Physiology*, *202*(1), 251-260.
- Hue, C. W. (2003). Number of characters a college student knows. *Chinese Journal of Psychology*, *31*, 300-339.
- Hue, C. W., & Tzeng, A. (2000). Lexical recognition task: A new method for the study of chinese character recognition. *Acta Psychologica Sinica*, *32*, 60-65.
- Ino, T., Nakai, R., Azuma, T., Kimura, T., & Fukuyama, H. (2009). Recognition and reading aloud of kana and kanji word: An fMRI study. *Brain Research Bulletin*, *78*(4-5), 232-239.

- James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: An fMRI study. *Brain*, *126*(Pt 11), 2463-2475.
- Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *Neuroimage*, *20*(2), 693-712.
- Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision*, *3*(7), 499-512.
- Jordan, K., Heinze, H. J., Lutz, K., Kanowski, M., & Jancke, L. (2001). Cortical activations during the mental rotation of different visual objects. *Neuroimage*, *13*(1), 143-152.
- Joseph, J. E., Partin, D. J., & Jones, K. M. (2002). Hypothesis testing for selective, differential, and conjoined brain activation. *Journal of Neuroscience Methods*, *118*(2), 129-140.
- Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679-685.
- Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Brain Research. Cognitive Brain Research*, *5*(1-2), 55-67.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302-4311.
- Kao, C. H., Chen, D. Y., & Chen, C. C. (2010). The inversion effect in visual word form processing. *Cortex*, *46*(2), 217-230.
- Kevan, A., & Pammer, K. (2009). Predicting early reading skills from pre-reading measures of dorsal stream functioning. *Neuropsychologia*, *47*(14), 3174-3181.

- Kohavi, R. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection*. Paper presented at the Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence.
- Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, 39(16), 2729-2737.
- Kourtzi, Z., & Huberle, E. (2005). Spatiotemporal characteristics of form analysis in the human visual cortex revealed by rapid event-related fMRI adaptation. *Neuroimage*, 28(2), 440-452.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, 20(9), 3310-3318.
- Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, 293(5534), 1506-1509.
- Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, 103(10), 3863-3868.
- Kuncheva, L. I., & Rodriguez, J. J. (2010). Classifier ensembles for fMRI data analysis: An experiment. *Magnetic Resonance Imaging*, 28(4), 583-593.
- Kuo, W. J., Yeh, T. C., Duann, J. R., Wu, Y. T., Ho, L. T., Hung, D., et al. (2001). A left-lateralized network for reading Chinese words: A 3T fMRI study. *Neuroreport*, 12(18), 3997-4001.
- Kuo, W. J., Yeh, T. C., Lee, C. Y., Wu, Y. T., Chou, C. C., Ho, L. T., et al. (2003). Frequency effects of Chinese character processing in the brain: An event-related fMRI study. *Neuroimage*, 18(3), 720-730.
- Kuo, W. J., Yeh, T. C., Lee, J. R., Chen, L. F., Lee, P. L., Chen, S. S., et al. (2004). Orthographic and phonological processing of Chinese characters: An fMRI study. *Neuroimage*, 21(4), 1721-1731.

- LaConte, S., Strother, S., Cherkassky, V., Anderson, J., & Hu, X. (2005). Support vector machines for temporal classification of block design fMRI data. *Neuroimage*, 26(2), 317-329.
- Larsson, J., & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience*, 26(51), 13128-13142.
- Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, 53(2), 513-536.
- Legge, G. E., Cheung, S. H., Yu, D., Chung, S. T., Lee, H. W., & Owens, D. P. (2007). The case for the visual span as a sensory bottleneck in reading. *Journal of Vision*, 7(2), 9 1-15.
- Legge, G. E., Mansfield, J. S., & Chung, S. T. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research*, 41(6), 725-743.
- Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision*, 2(2), 167-177.
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, 25(7), 963-977.
- Levi, D. M., Klein, S. A., & Chen, I. (2008). What limits performance in the amblyopic visual system: Seeing signals in noise with an amblyopic brain. *Journal of Vision*, 8(4), 1 1-23.
- Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience*, 4(5), 533-539.
- Liu, C., Zhang, W. T., Tang, Y. Y., Mai, X. Q., Chen, H. C., Tardif, T., et al. (2008). The Visual Word Form Area: Evidence from an fMRI study of implicit processing of Chinese characters. *Neuroimage*, 40(3), 1350-1361.

- Livingstone, M. S., Freeman, D. C., & Hubel, D. H. (1996). Visual responses in V1 of freely viewing monkeys. *Cold Spring Harbor Symposia on Quantitative Biology*, 61, 27-37.
- Lux, S., Marshall, J. C., Thimm, M., & Fink, G. R. (2008). Differential processing of hierarchical visual stimuli in young and older healthy adults: Implications for pathology. *Cortex*, 44(1), 21-28.
- Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Sciences*, 6(4), 176-184.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92(18), 8135-8139.
- Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, 5(1), 58-70.
- McKone, E. (2004). Isolating the special component of face recognition: Peripheral identification and a Mooney face. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1), 181-197.
- McKone, E., Brewer, J. L., MacPherson, S., Rhodes, G., & Hayward, W. G. (2007). Familiar other-race faces show normal holistic processing and are robust to perceptual stress. *Perception*, 36(2), 224-248.
- McKone, E., Martini, P., & Naykayama, K. (2003). Isolating holistic processing in faces (and perhaps objects) In M. A. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes*. (pp. 92-119). NY: Oxford University Press.
- Mechelli, A., Price, C. J., Noppeney, U., & Friston, K. J. (2003). A dynamic causal modeling study on category effects: Bottom-up or top-down mediation? *Journal of Cognitive Neuroscience*, 15(7), 925-934.
- Mosig, M. O., Lipkin, E., Khutoreskaya, G., Tchourzyna, E., Soller, M., & Friedmann, A. (2001). A whole genome scan for quantitative trait loci affecting milk protein percentage in Israeli-Holstein cattle, by means of selective milk DNA pooling

- in a daughter design, using an adjusted false discovery rate criterion. *Genetics*, 157(4), 1683-1698.
- Navon, D. (1977). Forest before trees - precedence of global features in visual-perception. *Cognitive Psychology*, 9(3), 353-383.
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424-430.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.
- Ostwald, D., Lam, J. M., Li, S., & Kourtzi, Z. (2008). Neural coding of global form in the human visual cortex. *Journal of Neurophysiology*, 99(5), 2456-2469.
- Pammer, K., Hansen, P., Holliday, I., & Cornelissen, P. (2006). Attentional shifting and the role of the dorsal pathway in visual word recognition. *Neuropsychologia*, 44(14), 2926-2936.
- Pammer, K., Lavis, R., Cooper, C., Hansen, P. C., & Cornelissen, P. L. (2005). Symbol-string sensitivity and adult performance in lexical decision. *Brain and Language*, 94(3), 278-296.
- Pammer, K., & Vidyasagar, T. R. (2005). Integration of the Visual and Auditory Networks in Dyslexia: A Theoretical Perspective. *Journal of Research in Reading*, 28, 320-331.
- Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, 46(28), 4646-4674.
- Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature*, 423(6941), 752-756.
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, 4(12), 1136-1169.

- Penny, W. D., Holmes, A. P., & Friston, K. J. (2003). Random effects analysis. . In R. S. J. Frackowiak (Ed.), *Human brain function, second edition*. (pp. 843-850). London: Academic Press.
- Perea, M., & Rosa, E. (2002). Does "whole-word shape" play a role in visual word recognition? *Perception and Psychophysics*, *64*(5), 785-794.
- Poldrack, R. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). The neural basis of visual skill learning: An fMRI study of mirror reading. *Cerebral Cortex*, *8*(1), 1-10.
- Polk, T. A., & Farah, M. J. (2002). Functional MRI evidence for an abstract, not perceptual, word-form area. *Journal of Experimental Psychology: General*, *131*(1), 65-72.
- Price, C. J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy*, *197 Pt 3*, 335-359.
- Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage*, *19*(3), 473-481.
- Proverbio, A. M., Wiedemann, F., Adorni, R., Rossi, V., Del Zotto, M., & Zani, A. (2007). Dissociating object familiarity from linguistic properties in mirror word reading. *Behavioral and Brain Functions*, *3*, 43.
- Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. New York: Prentice-Hall.
- Reinholz, J., & Pollmann, S. (2005). Differential activation of object-selective visual areas by passive viewing of pictures and words. *Brain Research. Cognitive Brain Research*, *24*(3), 702-714.
- Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, *47*(1), 25-57.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019-1025.

- Riesenhuber, M., & Poggio, T. (2002). Neural mechanisms of object recognition. *Current Opinion in Neurobiology*, *12*(2), 162-168.
- Riesenhuber, M., & Wolff, B. S. (2009). Task effects, performance levels, features, configurations, and holistic face processing: A reply to Rossion. *Acta Psychologica*, *132*(3), 286-292.
- Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, *126*, 2381-2395.
- Rossion, B., Dricot, L., Devolder, A., Bodart, J. M., Crommelinck, M., de Gelder, B., et al. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *12*(5), 793-802.
- Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral & Cognitive Neuroscience Reviews*, *1*(1), 63-75.
- Rumelhart, D. E., McClelland, J. L., & the PDP research group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume I*. Cambridge, MA: MIT Press.
- Ryan, L., & Schnyer, D. (2007). Regional specificity of format-specific priming effects in mirror word reading using functional magnetic resonance imaging. *Cerebral Cortex*, *17*(4), 982-992.
- Schendan, H. E., & Lucia, L. C. (2010). Object-sensitive activity reflects earlier perceptual and later cognitive processing of visual objects between 95 and 500ms. *Brain Research*, *1329*, 124-141.
- Schlaggar, B. L., Brown, T. T., Lugar, H. M., Visscher, K. M., Miezin, F. M., & Petersen, S. E. (2002). Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science*, *296*(5572), 1476-1479.
- Sekuler, E. B., & Behrmann, M. (1996). Perceptual cues in pure alexia. *Cognitive Neuropsychology*, *13*(7), 941-974.

- Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically- based intervention. *Biological Psychiatry*, *55*(9), 926-933.
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, *52*(2), 101-110.
- Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., et al. (2003). Neural systems for compensation and persistence: Young adult outcome of childhood reading disability. *Biological Psychiatry*, *54*(1), 25-33.
- Starrfelt, R., & Gerlach, C. (2007). The visual what for area: Words and pictures in the left fusiform gyrus. *Neuroimage*, *35*(1), 334-342.
- Stehling, M. K., Turner, R., & Mansfield, P. (1991). Echo-planar imaging: Magnetic resonance imaging in a fraction of a second. *Science*, *254*(5028), 43-50.
- Sun, F., Morita, M., & Stark, L. W. (1985). Comparative patterns of reading eye movement in Chinese and English. *Perception and Psychophysics*, *37*(6), 502-506.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical Publishers.
- Tan, L. H., Feng, C. M., Fox, P. T., & Gao, J. H. (2001). An fMRI study with written Chinese. *Neuroreport*, *12*(1), 83-88.
- Tan, L. H., Hoosain, R., & Peng, D. L. (1995). Role of presemantic phonological code in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 43-54.
- Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping*, *25*(1), 83-91.

- Tan, L. H., Liu, H. L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J. H. (2001). The neural system underlying Chinese logograph reading. *Neuroimage*, *13*(5), 836-846.
- Tan, L. H., Spinks, J. A., Gao, J. H., Liu, H. L., Perfetti, C. A., Xiong, J., et al. (2000). Brain activation in the processing of Chinese characters and words: A functional MRI study. *Human Brain Mapping*, *10*(1), 16-27.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *46*(2), 225-245.
- Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *Journal of Neurophysiology*, *98*(3), 1626-1633.
- Temple, E. (2002). Brain mechanisms in normal and dyslexic readers. *Current Opinion in Neurobiology*, *12*(2), 178-183.
- Teo, P. C., Sapiro, G., & Wandell, B. A. (1997). Creating connected representations of cortical gray matter for functional MRI visualization. *IEEE Transactions on Medical Imaging*, *16*(6), 852-863.
- Thomas, J. P. (1985a). Detection and identification: How are they related? *Journal of the Optical Society of America A. Optics and Image Science*, *2*(9), 1457-1467.
- Thomas, J. P. (1985b). Effect of static-noise and grating masks on detection and identification of grating targets. *Journal of the Optical Society of America A. Optics and Image Science*, *2*(9), 1586-1592.
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*(7), 1349-1357.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., et al. (1998). The retinotopy of visual spatial attention. *Neuron*, *21*(6), 1409-1422.

- Tsang, Y. K., & Chen, H. C. (2009). Do position-general radicals have a role to play in processing Chinese characters? *Language and Cognitive Processes*, 24(7/8), 947-966.
- Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., & Eden, G. F. (2003). Development of neural mechanisms for reading. *Nature Neuroscience*, 6(7), 767-773.
- Tyler, C. W., & Chen, C. C. (2000). Signal detection theory in the 2AFC paradigm: Attention, channel uncertainty and probability summation. *Vision Research*, 40(22), 3121-3144.
- Tyler, C. W., & Chen, C. C. (2006). Spatial summation of face information. *Journal of Vision*, 6(10), 1117-1125.
- Tyler, C. W., Likova, L. T., Chen, C. C., Kontsevich, L. L., Schira, M. M., & Wade, A. R. (2005). Extended concepts of occipital retinotopy. *Current Medical Imaging Reviews*, 1(3), 319-329.
- Tyler, C. W., & McBride, B. (1997). The Morphonome image psychophysics software and a calibrator for Macintosh systems. *Spatial Vision*, 10(4), 479-484.
- Tzeng, O. J., Hung, D. L., & Cotton, B. (1979). Visual interalisation effect in reading Chinese characters. *Nature*, 282(5738), 499-501.
- Vigneau, M., Jobard, G., Mazoyer, B., & Tzourio-Mazoyer, N. (2005). Word and non-word reading: What role for the Visual Word Form Area? *Neuroimage*, 27(3), 694-705.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron*, 55(1), 143-156.
- Wandell, B. A., Chial, S., & Backus, B. T. (2000). Visualization and measurement of the cortical surface. *Journal of Cognitive Neuroscience*, 12(5), 739-752.
- Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, 56(2), 366-383.

- Wang, M. Y., & Kuo, B. C. (2007). *The inversion effects for Chinese characters*. Paper presented at the Cognitive Neuroscience Society 2007 Annual Meeting.
- Wang, S. Y. (1973). The Chinese language. *Scientific American*, 228, 50-63.
- Wilson, H. R., & Wilkinson, F. (1998). Detection of global structure in Glass patterns: Implications for form vision. *Vision Research*, 38(19), 2933-2947.
- Yang, M. J., & Cheng, C. M. (1999). Hemisphere differences in accessing lexical knowledge of Chinese characters. *Laterality*, 4(2), 149-166.
- Yeh, S. L., & Li, J. L. (2002). Role of structure and component in judgments of visual similarity of Chinese characters. *Journal of Experimental Psychology: Human Perception and Performance*, 28(4), 933-947.
- Yeh, S. L., Li, J. L., Takeuchi, T., Sun, V., & Liu, W. R. (2003). The role of learning experience on the perceptual organization of Chinese characters. *Visual Cognition*, 10(6), 729-764.
- Yin, R. K. (1969). Looking at Upside-down Faces. *Journal of Experimental Psychology*, 81(1), 141-&.
- Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, 44(5), 889-898.
- Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology*, 15(24), 2256-2262.
- Zeki, S., Perry, R. J., & Bartels, A. (2003). The processing of kinetic contours in the brain. *Cerebral Cortex*, 13(2), 189-202.
- Zhang, J. Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2009). Legibility of Chinese characters in peripheral vision and the top-down influences on crowding. *Vision Research*, 49(1), 44-53.



Appendices

Appendix 1: Character stimuli in Chapter 2

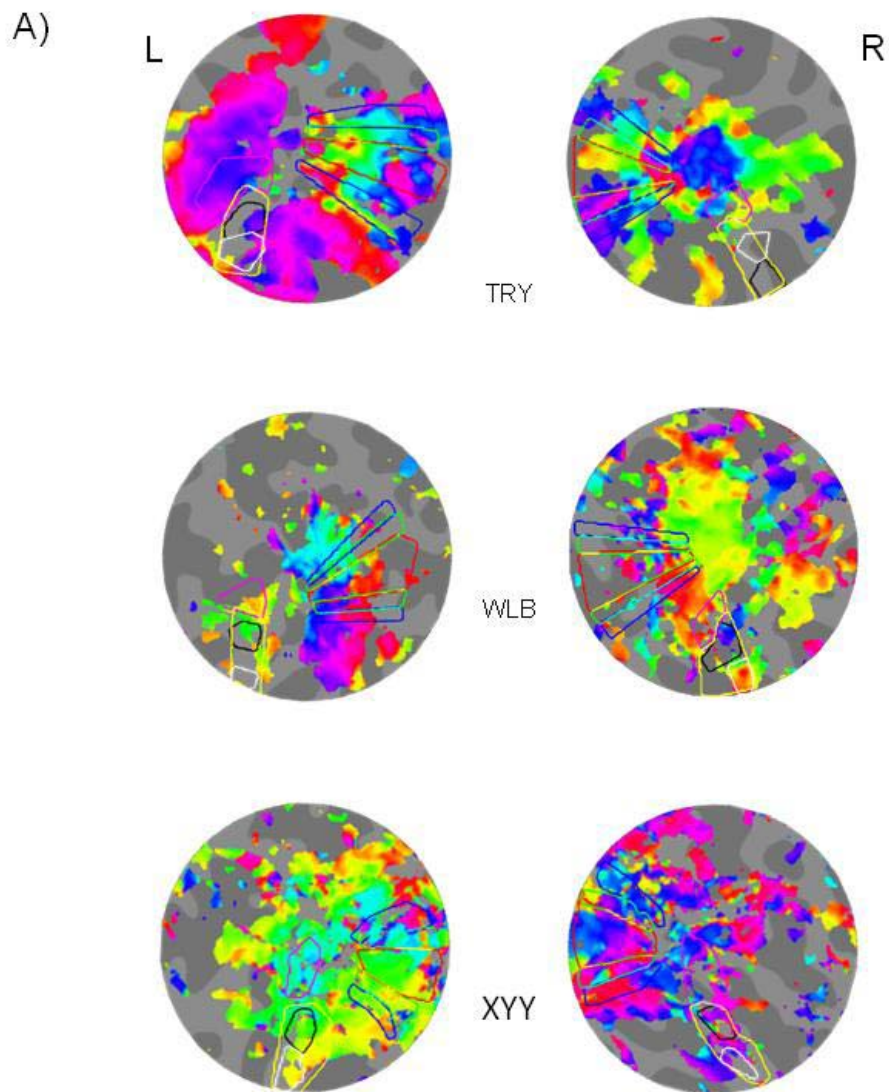
Real character	Pseudo-character	Non-character	<i>Jiagu</i> character
新	媯	齡	𠄎
說	媯	敦	𠄎
報	𧈧	𧈧	𧈧
期	鉞	𧈧	𧈧
情	緝	𧈧	𧈧
提	𧈧	𧈧	𧈧
設	𧈧	𧈧	𧈧
時	𧈧	𧈧	𧈧
就	𧈧	𧈧	𧈧
院	𧈧	𧈧	𧈧
都	𧈧	𧈧	𧈧
務	𧈧	𧈧	𧈧
被	𧈧	𧈧	𧈧
陳	𧈧	𧈧	𧈧
後	𧈧	𧈧	𧈧
經	𧈧	𧈧	𧈧
教	𧈧	𧈧	𧈧
認	𧈧	𧈧	𧈧

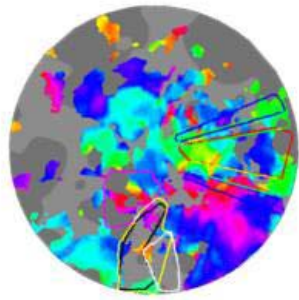
解	款	粘	𦉳
強	碇	𦉳	𦉳
程	殢	𦉳	𦉳
統	貶	𦉳	𦉳
張	獵	𦉳	𦉳
特	裨	𦉳	𦉳
海	砵	𦉳	𦉳
項	歛	𦉳	𦉳
隊	嶮	𦉳	𦉳
接	牀	𦉳	𦉳
除	猥	𦉳	𦉳
規	頽	𦉳	𦉳
將	鞞	𦉳	𦉳
結	肱	𦉳	𦉳
許	焮	𦉳	𦉳
請	倅	𦉳	𦉳
從	蝱	𦉳	𦉳
頭	紹	𦉳	𦉳
陸	雌	𦉳	𦉳
清	驥	𦉳	𦉳
線	飡	𦉳	𦉳
預	鱗	𦉳	𦉳
影	駟	𦉳	𦉳
校	忼	𦉳	𦉳

視	鞞	軫	𨾏
增	嫫	𨾏	𨾏
港	𨾏	𨾏	𨾏
雄	𨾏	𨾏	𨾏
餘	𨾏	𨾏	𨾏
條	𨾏	𨾏	𨾏
較	𨾏	𨾏	𨾏
標	𨾏	𨾏	𨾏
消	𨾏	𨾏	𨾏
能	𨾏	𨾏	𨾏
則	𨾏	𨾏	𨾏
理	𨾏	𨾏	𨾏
持	𨾏	𨾏	𨾏
施	𨾏	𨾏	𨾏
場	𨾏	𨾏	𨾏
現	𨾏	𨾏	𨾏
勁	𨾏	𨾏	𨾏
研	𨾏	𨾏	𨾏

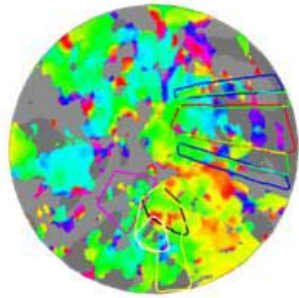
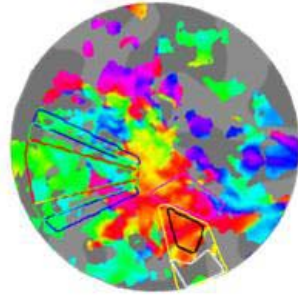
Appendix 2: Retinotopic activation for other 5

observers in Chapter 3

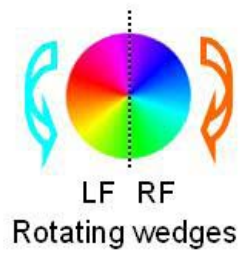
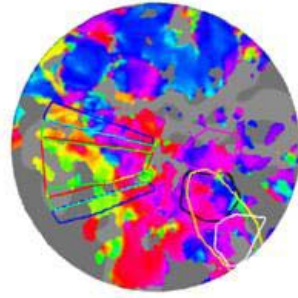




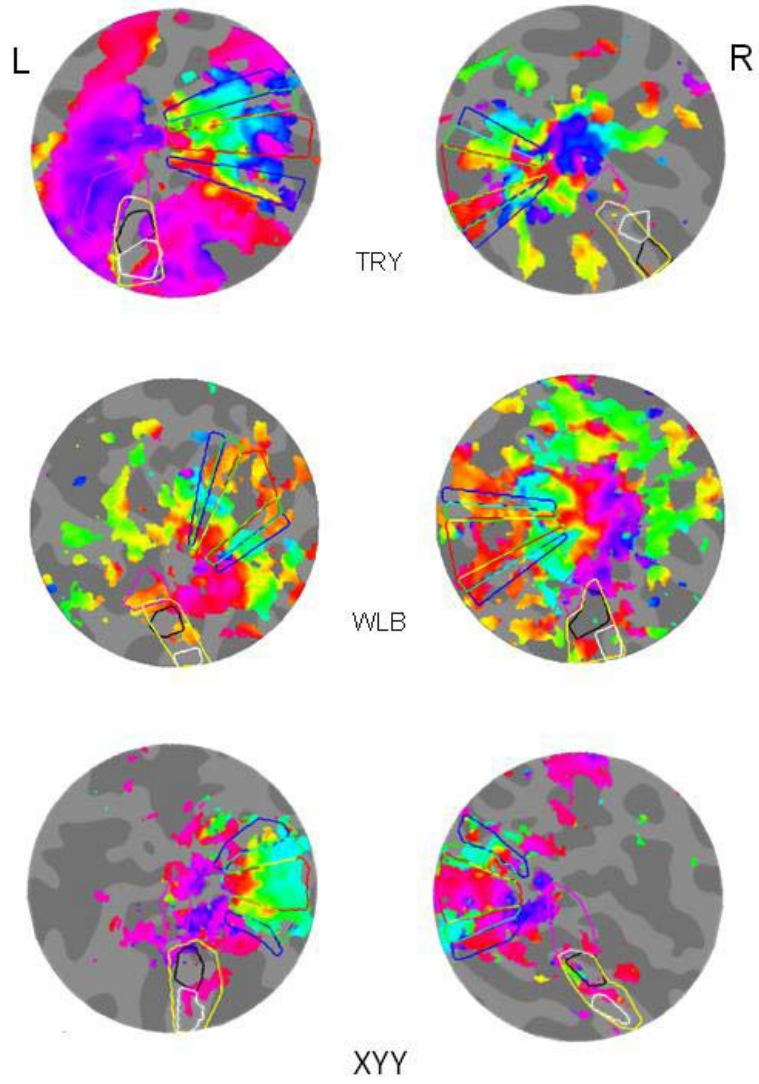
LYS



CKP



B)



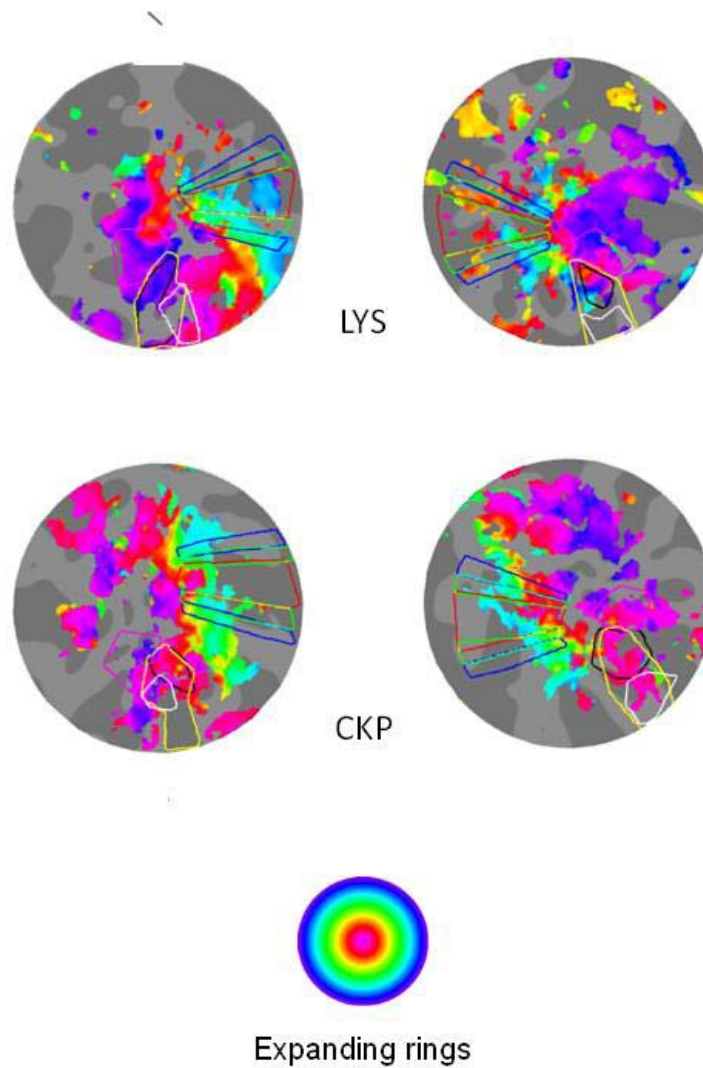
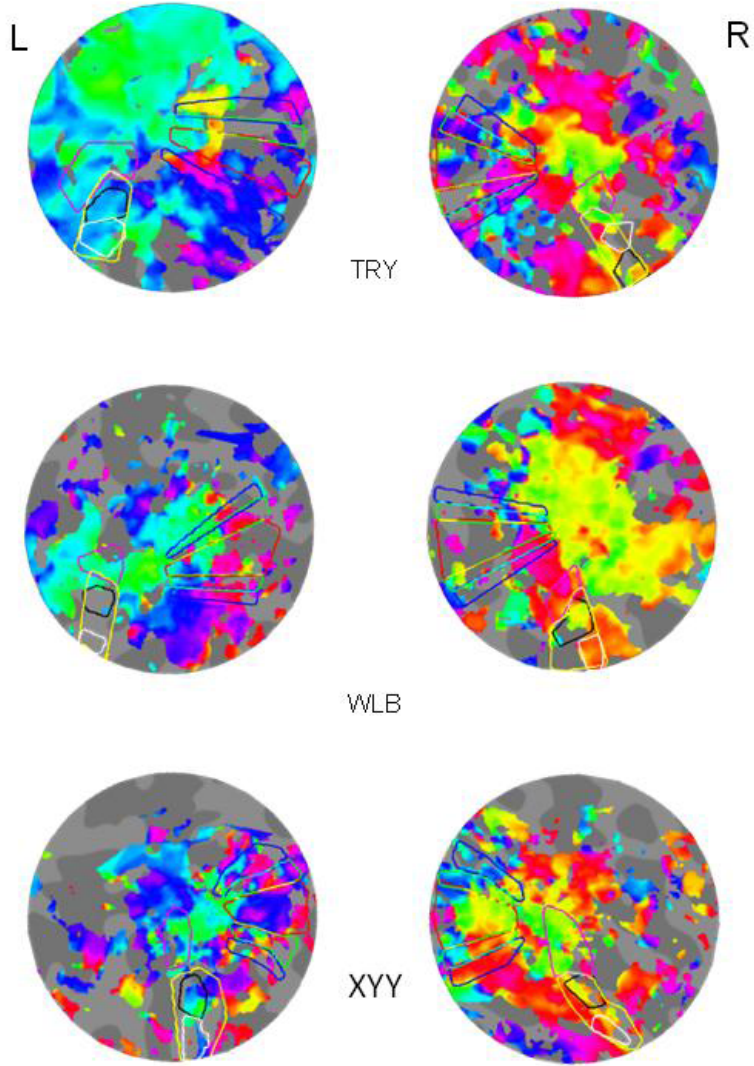
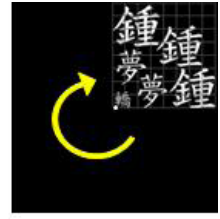
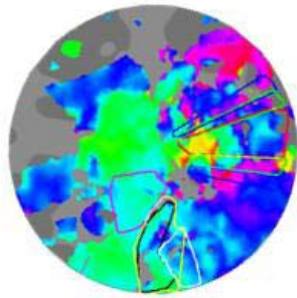
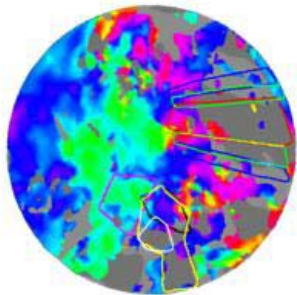
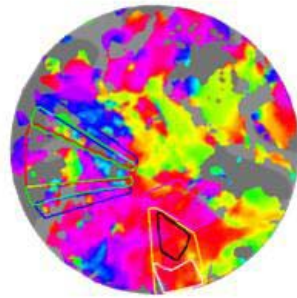


Figure A.1 Retinotopic activations for the A) rotating-wedge and B) expanding-ring checkerboards for other five observers. Radial color map segregates the left field(LF) and the right field(RF). Concentric color map separates the fovea and periphery. The outlines of red, green and blue denote V1, V2, and V3, respectively. LH (left hemisphere), RH (right hemisphere)

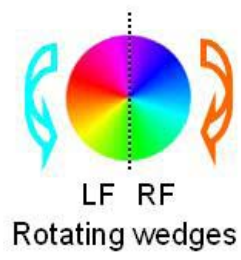
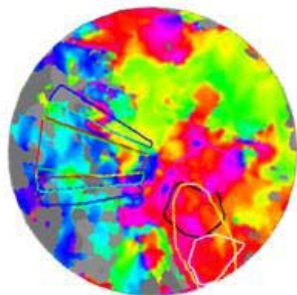


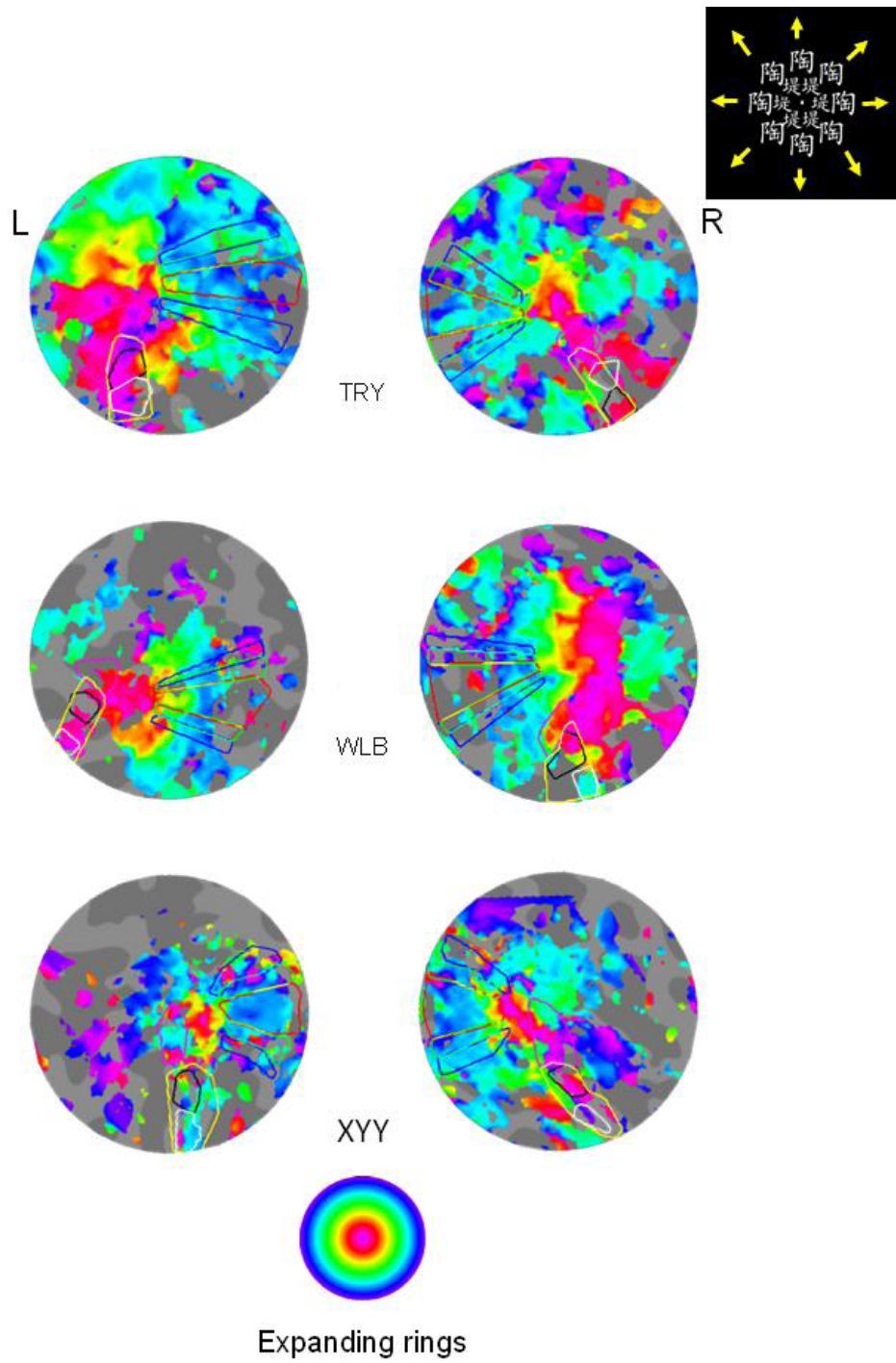


LYS



CKP





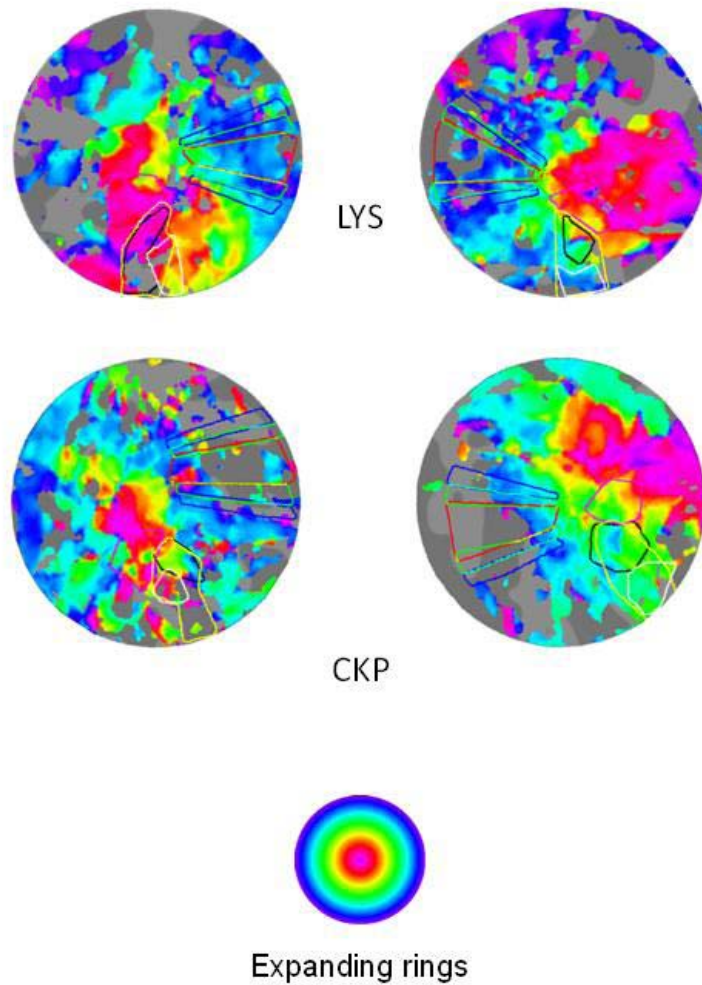
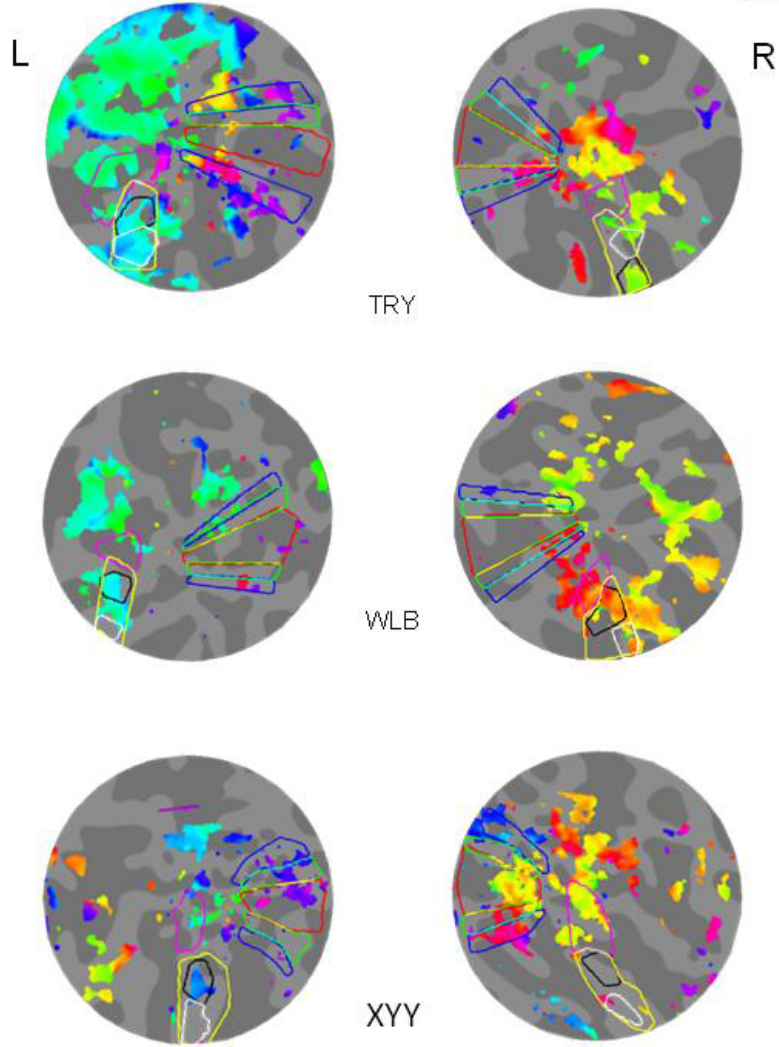
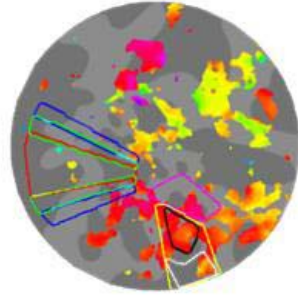
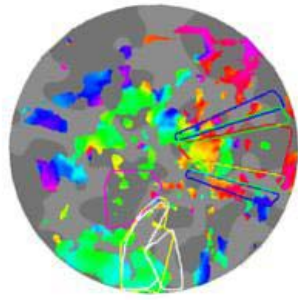
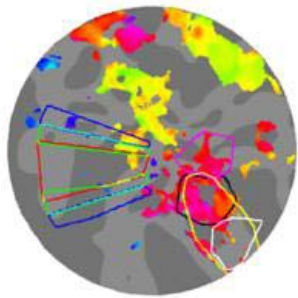
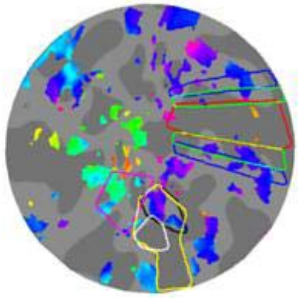


Figure A.2 Retinotopic activations for the A) rotating-wedge and B) expanding-ring characters with a blank background for other five observers. Radial color map segregates the left field(LF) and the right field(RF). The magenta, yellow, black and white borders denote the LO, FG, VWFA, and FFA, respectively. LH (left hemisphere), RH (right hemisphere)

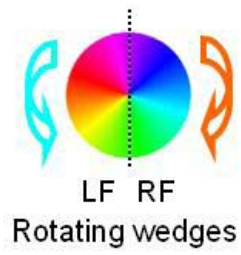


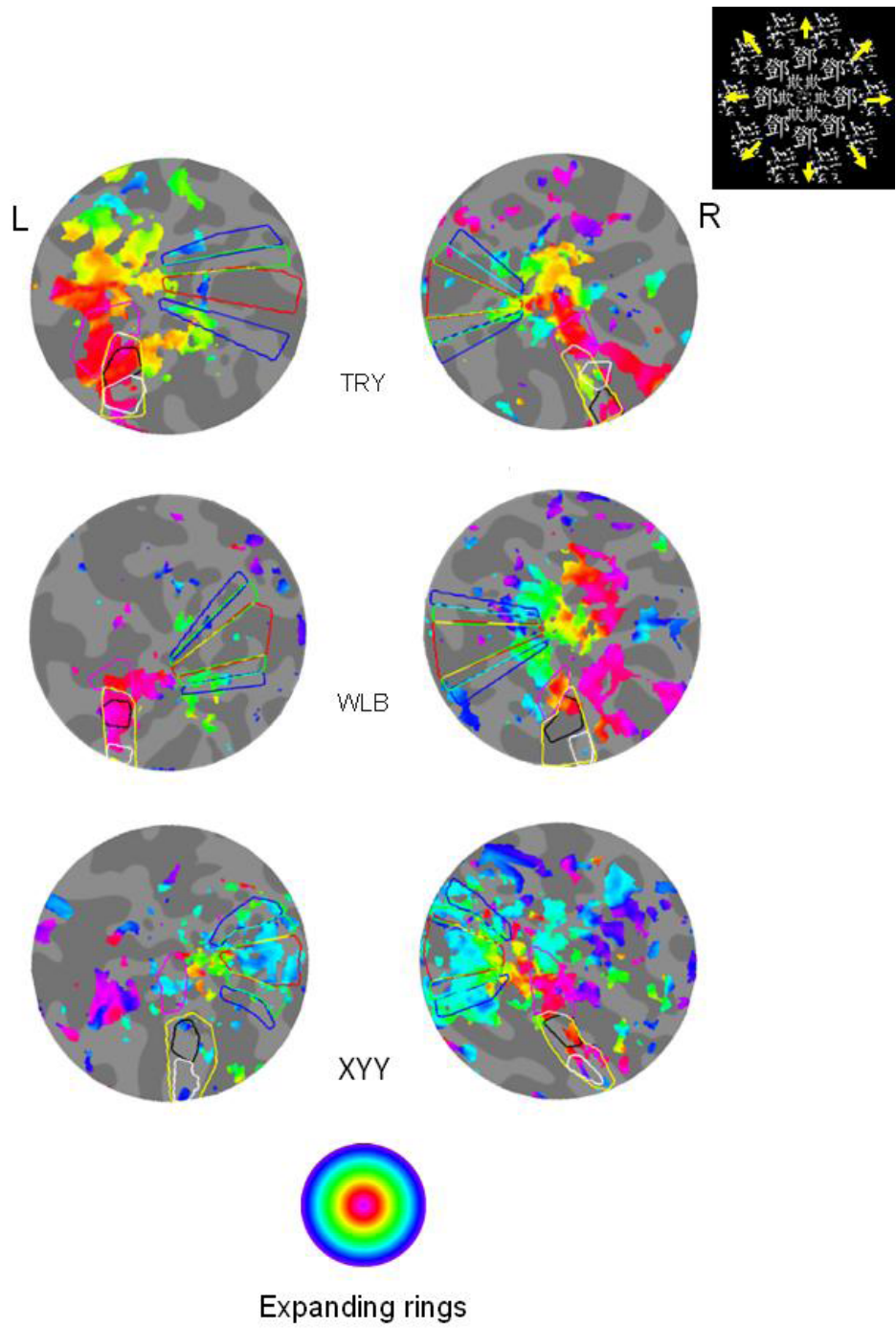


LYS



CKP





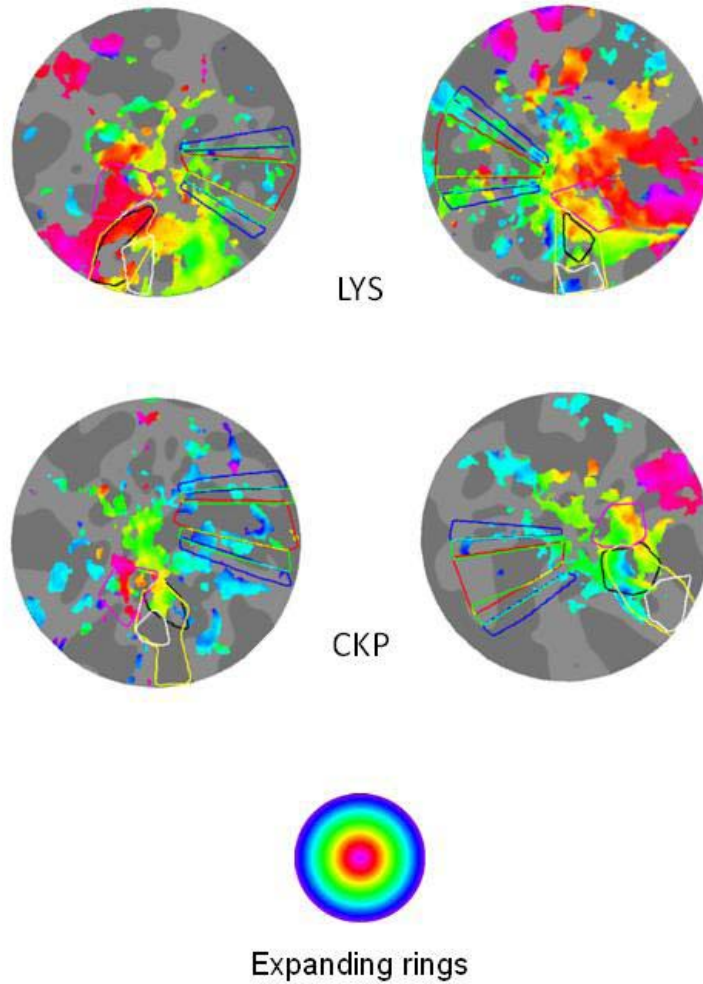


Figure A.3 Retinotopic activations for rotating-wedge and expanding-ring characters with a scrambled background for other five observers. Radial color map segregates the LF and the RF. Concentric color map separates the fovea and periphery. V1 to V3 are defined by the rotating-wedge and expanding-ring checkerboards. The magenta, yellow, black and white borders denote the LO, FG, VWFA, and FFA, respectively. LH (left hemisphere), RH (right hemisphere)

Appendix 3: Character stimuli in Chapter 4 and 5

Real character	Non-character	Lexical components	Non-lexical components	<i>Jiagu</i> character
新	𪛗	寺	叟	𪛗
說	𪛗	見	易	𪛗
報	𪛗	幸	亲	𪛗
期	𪛗	是	务	𪛗
情	𪛗	黑	息	𪛗
提	𪛗	𪛗	𪛗	𪛗
設	𪛗	𪛗	𪛗	𪛗
時	𪛗	長	𪛗	𪛗
就	𪛗	重	𪛗	𪛗
院	𪛗	青	壘	𪛗
都	𪛗	泉	𪛗	𪛗
務	𪛗	登	惠	𪛗
被	𪛗	賣	𪛗	𪛗
陳	𪛗	票	𪛗	𪛗
後	𪛗	祭	𪛗	𪛗
經	𪛗	專	段	𪛗
教	𪛗	襄	𪛗	𪛗
認	𪛗	容	𪛗	𪛗
解	𪛗	𪛗	𪛗	𪛗
強	𪛗	竟	𪛗	𪛗

程	咎	責	甬	𠵼
統	軻	聿	敦	𠵼
張	𠵼	帝	封	𠵼
特	𠵼	章	𠵼	𠵼
海	𠵼	喬	𠵼	𠵼
項	𠵼	豐	𠵼	𠵼
隊	𠵼	卑	扁	𠵼
接	𠵼	甫	散	𠵼
除	𠵼	韋	𠵼	𠵼
規	𠵼	賞	𠵼	𠵼
將	𠵼	俞	𠵼	𠵼
結	𠵼	昏	矣	𠵼
許	𠵼	革	呈	𠵼
請	𠵼	亨	𠵼	𠵼
從	𠵼	帛	𠵼	𠵼
頭	𠵼	每	𠵼	𠵼
陸	𠵼	某	𠵼	𠵼
清	𠵼	盾	𠵼	𠵼
線	𠵼	昆	矣	𠵼
預	𠵼	唐	𠵼	𠵼
影	𠵼	甚	𠵼	𠵼
校	𠵼	爭	𠵼	𠵼
視	𠵼	翟	𠵼	𠵼
增	𠵼	鬼	𠵼	𠵼

港	𧈧	卓	𧈧	𧈧
雄	𧈧	叟	豕	𧈧
餘	𧈧	掌	耳	𧈧
條	𧈧	豈	盜	𧈧
較	𧈧	甬	至	𧈧
標	𧈧	賁	夕	𧈧
消	𧈧	秀	盍	𧈧
能	𧈧	亶	皋	𧈧
則	𧈧	奐	卮	𧈧
理	𧈧	表	菑	𧈧
持	𧈧	危	盃	𧈧
施	𧈧	君	蕇	𧈧
場	𧈧	星	𧈧	𧈧
現	𧈧	旁	𧈧	𧈧
勁	𧈧	春	育	𧈧
研	𧈧	亥	展	𧈧

Curriculum Vitae

Education:

2011, National Taiwan University, Ph. D. in Psychology

2001, National Cheng-Chi University, M.S. in Psychology

1998, National Cheng-Chi University, B.S. in Psychology

Research Interests:

Vision, Object perception, Language processing, Neuroimaging

Publications:

Articles

Kao, C.H., Chen, D.Y., Chen, C.C. (2010). The inversion effect in visual word form processing. *Cortex*, 46, 217-230.

Kao, C.H., Hue, C.W., Tseng, Y.H., Lo, M. (2009). Detection errors in proofreading Chinese texts: A study of Chinese reading units and word inferiority effect in detecting character-position transposed words. *Chinese Journal of Psychology*, 51, 21-36.

Kao, C.H., Hue, C.W., Lo, M. (2006). Indigenous factors for studies on Chinese speaking aphasics. *Research in Applied Psychology*, 29, 13-15.

Hue, C.W., Lo, M., **Kao, C.H.** (2006). Neurologically-oriented cognitive psychology : Are we ready for it? *Research in Applied Psychology*, 29, 9-11.

Hue, C.W., **Kao, C.H.**, & Lo, M. (2005). Association norms for 600 Chinese characters. *Monographs of Chinese Journal of Psychology* (Serial No. 0501).

Conference Presentations

Kao, C.H., & Chen, C.C. (2010). The role of local and global configuration for the visual word form processing in the occipitotemporal cortex. Poster presented at 40th Society for Neuroscience Annual Meeting, San Diego, CA.

Kao, C.H., & Chen, C.C. (2010). Inversion effect in visual word forms: the role of spatial configurations and character components. Paper presented at 2010 Asia-Pacific Conference on Vision, Taipei, Taiwan. (Slide session)

Kao, C.H., Hue, C.W., Lu, C.C., Tseng, Y.H., Chen, Y.F., & Kou, Y.W. (2009). Factors Affecting Chinese Compound Word Representation and Semantic Priming Effect. Paper presented at the 19th International Symposium on Theoretical & Applied Linguistics, Thessaloniki, Greece. (Slide session)

Kao, C.H., Hue, C.W., Lu, C.C., Tseng, Y.H., Chen, Y.F., & Kou, Y.W. (2009). Factors Affecting Chinese Compound Word Representation and Semantic Priming Effect. Paper presented at the 19th International Symposium on Theoretical & Applied Linguistics, Thessaloniki, Greece. (Slide session)

Kao, C.H., Hue, C.W., Tseng, Y.H., Lo, M., Tsai, F.C. (2007). Detecting Character Position Transposed Chinese Words: A Study of Chinese version of Word Inferiority Effect. Paper presented at the 7th Biennial conference Asian association of social psychology, Sabah, Malaysia. (Slide session)

- Kao, C.H.**, Chen, C.C. (2007). The detection and discrimination of linguistic stimuli in the foveal and the peripheral vision. Poster presented at the 30th European Conference on Visual Perception, Arezzo, Italy.
- Kao, C.H.**, Chen, D.Y., Hue, C.W., Chen, J.H., Chen, C.C. (2006). Character inversion effect in the human occipitotemporal regions: an fMRI study. Poster presented at 26th Society for Neuroscience Annual Meeting, Atlanta, GA
- Kao, C.H.**, Chen, C.C., Chen, D.Y., Hue, C.W., Chen, J.H. (2006). Retinotopic representation in the fusiform gyrus: Evidence from the Chinese character processing. Workshop on Brain and Cognition, Taipei, Taiwan. (Slide session)
- Kao, C.H.**, Hue, C.W., Tsai, F.C., Lo, M., Chen, Y.F. (2005). Word Inferiority Effect in Detecting Position Transposed Character Pairs. Paper presented at the Processing Chinese and Other East Asian Languages, Hong Kong, China. (Slide session)
- Kao, C.H.**, Chen, C.C., Chen, D.Y., Hue, C.W., Chen, J.H. (2005) The retinotopy of ventral occipital complex: Evidence of Chinese characters. Poster presented at 25th Society for Neuroscience Annual Meeting, Washington D. C.
- Kao, C.H.**, Chen, C.C., Chen, D.Y., Hue, C.W., Chen, J.H. (2005). The retinotopy of ventral occipital temporal complex with Chinese characters. International symposium on cerebral cortical organization and function, Taipei, Taiwan. (Slide session)
- Kao, C.H.**, Hue, C.W., Chen, C.C., Chen, D.Y., Chen, J.H. (2004). The role of the left fusiform gyrus in the processing of Chinese characters. Paper

presented at 24th Society for Neuroscience Annual Meeting, San Diego,
CA. (Slide session)

