Vision Research 49 (2009) 44–53

Contents lists available at ScienceDirect

Vision Research

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

Legibility of Chinese characters in peripheral vision and the top-down influences on crowding

Jun-Yun Zhang [a,1], Ting Zhang [a,1], Feng Xue [b], Lei Liu [c,*], Cong Yu [a,*]

[a] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China
[b] EENT Hospital, Fudan University, Shanghai, China
[c] School of Optometry, University of Alabama at Birmingham, Birmingham, AL, USA

ARTICLE INFO

Article history:
Received 5 May 2008
Received in revised form 24 September 2008

Keywords:
Chinese characters
Legibility
Crowding
Top-down influences
Peripheral vision

ABSTRACT

Written Chinese is distinct from alphabetic languages because of its enormous number of characters with a great range of spatial complexities (stroke numbers). In this study we investigated the impact of spatial complexity on legibility of Chinese characters as well as associated crowding in peripheral vision. Our results showed that for isolated characters, threshold sizes of complex characters increased faster with retinal eccentricity than did those of simple characters, suggesting possible "within-character" crowding among parts of complex Chinese characters. However, such "within-character" crowding was rendered negligible by strong "between-character" crowding introduced by flankers. When the target and flankers belonged to different complexity groups, the intensity and extent of crowding were greatly reduced, which could be explained by top-down influences as well as lower-level mechanisms. We suggest that crowding can be attributed to multiple mechanisms at different levels of visual processing.

1. Introduction

Most studies of letter legibility use Roman letters. Roman letters are highly stylish visual stimuli that are made of a small number of strokes, have no discernible parts, and are relatively uniform in spatial complexity as a stimulus set. It is less clear how much of our knowledge obtained from such stimuli can be applied to legibility of Chinese characters (CCs) that contain 1 to as many as 52 strokes, and thus have a wide range of spatial complexities. Recently we reported a study on legibility of CCs in foveal vision (Zhang, Zhang, Xue, Liu, & Yu, 2007), in which we measured threshold (acuity) sizes for six groups of frequently used CCs from low to high spatial complexities, and determined the relationship between legibility and optical defocus for Landolt C, Snellen E and three groups of CCs representing low, medium, and high spatial complexities. Our results showed that CC acuity sizes increase steadily with stimulus complexity, though at a slower rate than what would be expected if visual acuity is based on discerning the finest details of the stimuli. Moreover, the acuity size vs. optical defocus functions of the three CCs groups and Snellen E have similar slopes, differing only by a vertical shift (approximately one, two, and three lines above E acuity on an acuity chart, respectively), suggesting the feasibility of using Snellen E acuity, which is the current standard optotype for acuity testing in China, to derive the legibility of CCs in foveal vision. To understand the slower rate of acuity size increase against spatial complexity, we also developed a geometric moment model, in which we propose that human letter recognition performance near the acuity limit can be accounted for by a set of global features described by easy-to-visualize and perceptually meaningful low-order geometric moments (i.e., the ink area, variance, skewness, and kurtosis; manuscript under review).

The current study extends our work to the legibility of CCs, as well as crowding, in peripheral vision. We are particularly interested in two distinct characteristics of CCs that could affect peripheral character legibility and crowding in ways not normally evident when alphabetic stimuli are used. First, the majority of CCs are spatially complicated. Only 4% of CCs are single-body characters (e.g., 个) that have a small number of strokes, and cannot be divided graphically into smaller and meaningful parts. The rest are compounds (e.g., 需) that are made of multiple parts, each of which is a single-body character or its derivation, and are packed in the same square area as the single-body CCs. We suspect that interactions among these parts could interfere with recognition of a complex CC as a whole, and such interactions, or "within-character" crowding, could be magnified in the periphery. If this is indeed the case, acuities of different complexity CC groups may have different spatial scaling functions in the periphery, and thus may not be derived from a standard measurement like E acuity, as we showed previously for foveal vision (Zhang et al., 2007), without

* Corresponding authors.
E-mail addresses: liul7788@uab.edu (L. Liu), yucong@bnu.edu.cn (C. Yu).
[1] These authors contributed equally to this work.

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.09.021

proper compensations of scaling differences among CC groups. Such a possibility would have important clinical implications in evaluating peripheral vision of patients who read text that contain characters of different spatial complexities.

To address this issue, in the first part of the study, we measured threshold sizes of single CCs of various complexities at different retinal eccentricities. By comparing the slopes of spatial scaling functions for different complexity CCs groups, we revealed an inferiority of complex CCs to simple CCs in the visual periphery, possibly indicating "within-character" crowding among parts of complex CCs. We also measured threshold sizes of flanked CCs in a trigram configuration to assess the impact of "within-character" crowding on regular "between-character" crowding.

The second distinct characteristic of CCs we are particularly interested in is that, in real-world Chinese text, more than often is a character flanked by characters of different spatial complexities. Such configurations are rarely seen in alphabetic languages because alphabetic letters tend to have similar spatial complexities. In cases where the target and flanking characters have different spatial complexities, some basic stimulus properties, such as the brightness and the spatial frequency contents, are different between the target and flankers. These and other physical stimulus differences including shape, size, polarity, etc., are known to affect crowding by segregating the target and flankers (Chung, Levi, & Legge, 2001; Hess, Dakin, & Kapoor, 2000; Kooi, Toet, Tripathy, & Levi, 1994; Nazir, 1992). Moreover, a Chinese reader knows naturally that the target and flanking characters with very different spatial complexities in a trigram configuration, such as 个需十, are drawn from different stimulus groups, so that he or she will not report a flanking character as the target. There is evidence that such misreporting contributes to crowding (Strasburger, 2005). Therefore, both stimulus differences and high-level top-down influences may affect crowding when the target and flanking characters differ in complexity.

In the second part of this study we assessed the impact of target–flanker complexity contrast on crowding. We also designed experiments to isolate the top-down influence on crowding, using not only CCs but also English Sloan letters. Moreover, after isolation of top-down influences, we were able to manipulate stimulus physical features to identify lower-level mechanisms underlying crowding. On the basis of our results, as well as previously reported findings, we propose an eclectic view that uses multiple mechanisms at multiple processing levels to explain crowding.

2. Methods

2.1. Observers and apparatus

Six observers with normal or corrected-to-normal vision participated in the study. All observers were young (mean age = 23.3 years) native Chinese speakers with college education and at least 6 years of training in reading and writing English. Observers ZJ and ZT were coauthors and were experienced in psychophysical experiments. The others were new to psychophysical observations and were unaware of the purposes of the study. Written informed consent was obtained from all observers prior to the tests.

The stimuli were generated by a Matlab-based WinVis program (Neurometrics Institute, Oakland, CA) and were presented on a 21-in. Sony G520 color monitor (2048 pixel $\times$ 1536 pixel, 0.189 mm $\times$ 0.189 mm per pixel, 75 Hz frame rate). The minimal and maximal luminance of the monitor was 1.18 and 91 cd/m^2, respectively. Viewing was monocular in a dimly lit room. A head-and-chin rest was used to stabilize the head position.

2.2. Stimuli

The test stimuli (Fig. 1a) consisted of one group of English Sloan letters and four groups of CCs. Each stimulus group contained five letters or characters (with exception in Experiment IV) with similar legibility as determined in a previous study (Zhang et al., 2007). In that study, 500 most frequently used CCs were selected and categorized into six groups according to the number of strokes (CC1–CC6 groups, from 2–4 to 16–18 strokes/character). Then ten characters were selected from each group based on intermediate Euclidean distances of character bitmaps, pronunciation, and spatial configuration. The legibility of these characters, along with ten Sloan letters, was measured in young normal observers using a rigorous psychophysical method. Based on these measurements, five stimuli with the most similar legibility within each group were selected for the use in the current experiments (CC2 and CC5 were not used). Since this article is part of a series of studies of Chinese character acuity, recognition and reading, which uses some or all six CC groups of different complexities, we chose to use these group names to be consistent with other articles. The bitmaps of the Sloan letters and CCs had the same width and height (50 $\times$ 50 pixels). The Sloan letter had uniform stroke width equivalent to 1/5 of the letter height. Font type bold Heiti (black font) was used for CCs because the strokes had relatively uniform width and were free of serif. To fit different number of strokes into the same area, stroke widths became gradually thinner as the characters became more complex. For the 50 $\times$ 50 pixel bitmaps we used, vertical stroke widths shifted from predominantly 7 pixels in CC1 to predominantly 6 pixels in CC6, and horizontal stroke widths shifted from 5–6 pixels in CC1 to 4–5 pixels in CC6.

The spatial complexity of the stimuli was also described by stroke frequency (Zhang et al., 2007). Each letter or character was sliced at 6 direction/position combinations: horizontal on the upper and lower halves, vertically on the left and right halves, and obliquely at 45° and 135° on the central portion of the stimuli. From each slicing we obtained the average crossed strokes and calculated the maximum of the 6 slicing as the stroke frequency. The average stroke frequency for the Sloan letters was 2.0 strokes/letter. The average stroke frequencies for the six groups of CCs increased monotonically from 2.2 to 5.5 strokes/character (Zhang et al., 2007).

2.3. Procedure

The target was a black Sloan letter or Chinese character presented on a full-screen full-luminance white background. The target was presented either alone or was flanked by two horizontally aligned letters or characters (trigram). The target could be any member of a stimulus group, and the two flankers were always

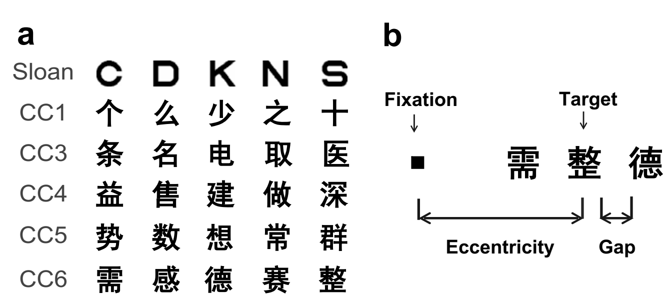

Fig. 1. Stimuli. (a) Sloan letters and four groups of CCs with various spatial complexities used in the study. CC1: 2–4 strokes; CC3: 8–9 strokes; CC4: 11–12 strokes; CC6: 16–18 strokes. (b) Stimulus configuration for peripheral testing when flankers were present. The flanking characters had the same size as the target, and the edge-to-edge gap between target and flanker was one character wide.