Vision Research 97 (2014) 45–51

Contents lists available at ScienceDirect

Vision Research

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

A crowdful of letters: Disentangling the role of similarity, eccentricity and spatial frequencies in letter crowding

Sacha Zahabi, Martin Arguin *

Centre de recherche en Neuropsychologie et Cognition, Département de psychologie, Université de Montréal, Canada

ARTICLE INFO

Article history:
Received 28 March 2013
Received in revised form 5 February 2014
Available online 19 February 2014

Keywords:
Visual
Letter identification
Spatial frequencies
Similarity

ABSTRACT

The present study investigated the joint impact of target–flanker similarity and of spatial frequency content on the crowding effect in letter identification. We presented spatial frequency filtered letters to neurologically intact non-dyslexic readers while manipulating target–flanker distance, target eccentricity and target–flanker confusability (letter similarity metric based on published letter confusion matrices). The results show that high target–flanker confusability magnifies crowding. They also reveal an intricate pattern of interactions of the spatial frequency content of the stimuli with target eccentricity, flanker distance and similarity. The findings are congruent with the notion that crowding results from the inappropriate pooling of target and flanker features and that this integration is more likely to match a response template at a subsequent decision stage with similar than dissimilar flankers. In addition, the evidence suggests that crowding from similar flankers is biased towards relatively high spatial frequencies and that crowding shifts towards lower spatial frequencies as target eccentricity is increased.

1. Introduction

Visual crowding refers to the difficulty of accurately identifying a peripheral visual stimulus when it is flanked by other items. The currently accepted account of crowding assumes that target features are normally detected independently from flanker features, provided that the distance between them is sufficiently large (Levi, 2008). However, when the target–flanker distance is too short, features from both items fall within the same integration fields. Target and flanker features then become difficult to segregate, which interferes with target identification (Pelli, Palomares, & Majaj, 2004). Given that integration fields increase in size as one goes from the fovea to the visual periphery, eccentric targets are more susceptible to crowding with reduced target–flanker distances. Congruently with this account, Levi (2008) proposed a two-stage model of visual feature processing involving first the detection of simple features (in V1), followed by their integration (beyond V1).

The results from a considerable number of studies have identified three major factors that determine crowding. Thus, the magnitude of crowding is a function of inter-stimulus distance and there is a critical spacing beyond which crowding no longer occurs (Pelli, Palomares, & Majaj, 2004). Also, this critical spacing is directly proportional to eccentricity (Pelli, Palomares, & Majaj, 2004; see also Bouma, 1970). Finally, the more similar the flankers are to the target, the more they affect its identification (e.g. Bernard & Chung, 2011; Chung, Levi, & Legge, 2001; Estes, 1982; Freeman, Chakravarthi, & Pelli, 2012; Hess, Dakin, & Kapoor, 2000a; Kooi et al., 1994; Poder, 2007; Shapiro & Krueger, 1983). For instance, Kooi et al. (1994) have demonstrated, using a task requiring observers to identify the orientation of a T flanked by three other T's, that target–flanker dissimilarity in terms of contrast polarity, depth or orientation improved identification performance (see also Hess et al., 2000a). In the letter recognition domain, Bernard and Chung (2011) have shown that the error rates in the identification of a flanked target letter increase with the shape similarity of flankers (see also Estes, 1982; Krumhansl & Thomas, 1977). Relatedly, Freeman, Chakravarthi, and Pelli (2012) have demonstrated that when an error is made in the identification of a flanked letter, similar flankers are much more likely to be reported than dissimilar flankers.

Letters contain a wide range of spatial frequencies and many recent studies have attempted to determine the range of spatial frequencies that are preferentially used by the visual system to identify letters (Grainger, Rey, & Dufau, 2008). This question has profound implications given that our ability to read a word depends first and foremost on the efficiency of our visual system to identify each letter (Pelli, Farell, & Moore, 2003). Critical-band masking studies have shown that visual noise around 3 cycles/let-

* Corresponding author. Address: Département de psychologie, Pavillon Marie-Victorin, Université de Montréal, C.P. 6128, Succ. Centre-ville, Montréal, Qc H3C 3J7, Canada. Fax: +1 514 343 5787.

E-mail address: martin.arguin@umontreal.ca (M. Arguin).

http://dx.doi.org/10.1016/j.visres.2014.02.001
0042-6989/

ter impairs letter identification performance the most (Majaj et al., 2002; Solomon & Pelli, 1994). This suggests that the optimal spatial frequencies for letter identification are around 3 cycles/letter. Congruent findings were obtained through the contrast thresholds for the identification of band-pass filtered letters (Chung, Legge, & Tjan, 2002). In further support, a study by Fiset et al. (2008), using the Bubbles technique (Gosselin & Schyns, 2001), has revealed that spatial frequencies between 2 and 4 cycles/letter provide the most useful information for letter identification. According to Chung, Legge, and Tjan (2002), these optimal spatial frequencies are determined by the intersection of the contrast sensitivity function of human vision with the spatial frequency content of the stimuli that best discriminates among the letters of the alphabet. Relatedly, an important feature of the range of spatial frequencies that dominate letter recognition is that it shifts towards lower retinal frequencies with increasing eccentricity (Chung, Legge, & Tjan, 2002).

Grainger, Rey, and Dufau (2008) point out that more information useful for letter identification is available in high-pass filtered letters than low-pass filtered ones (Chung, Legge, & Tjan, 2002; Parish & Sperling, 1991). Congruently, low spatial frequencies seem to exacerbate the difficulty in discriminating among visually similar letters; i.e. the letter confusability effect. The confusability value for a particular letter is determined from the error rates of normal observers in a task of single letter identification using very brief displays[1] (see Fiset et al., 2008; for a brief review). With words made of letters with a high confusability value, the word recognition performance of normal readers is significantly deteriorated relative to low confusability content with low-pass stimuli (Fiset, Arguin, & Fiset, 2006; Fiset et al., 2006). In contrast, normal readers are impervious to the effect of letter confusability with normal print or with high-pass or broadband filtered letters. Relatedly, an apparent bias towards low spatial frequencies seems implicated in the particular susceptibility of letter-by-letter dyslexics to the letter confusability effect in their word recognition performance with normal print (Arguin, Fiset, & Bub, 2002; Fiset et al., 2005, 2006).

Few studies have examined the role of spatial frequencies in visual crowding. Hess and his collaborators (Hess et al., 2000a; Hess, Dakin, Kapoor & Tewfik, 2000b) reported that the most relevant spatial frequencies for visual processing are shifted towards higher values under crowded conditions. At the fovea, this effect is entirely explained by a shift in the power spectra of the stimulus but this is not the case in the periphery (beyond 5 deg eccentricity), where an alteration of visual processing must be assumed. Chung and Tjan (2007) presented normal observers with spatial frequency filtered target letters flanked on either side by other letters, with three different levels of spacing. Similarly to Hess et al. (2000a, 2000b), their results show that the visual system slightly shifts its sensitivity to higher spatial frequencies when the target letter is surrounded by flankers, but this effect only occurred at the shortest flanker distance (i.e. 0.8*x*; *x* being the height of letter x for the particular font used, a standard metric in the literature on crowding). They also report that this shift cannot solely be accounted by an alteration of the physical properties of the stimuli, whether they are displayed at the fovea or at 5 deg eccentricity.

Chung, Levi, and Legge (2001) have also manipulated the physical properties of visual stimuli to examine the crowding effect in normal readers. Spatial frequency filtered letters were presented with or without flankers at the fovea or at 5° eccentricity. The dependent variable was the contrast threshold required to identify the target letter. The results showed that shorter flanker distances produce a contrast threshold elevation peak when the spatial frequency content of the flankers is similar to that of the target and that this threshold elevation diminishes with a reduction of spatial frequency similarity. This effect was qualitatively the same at the fovea and at 5° eccentricity.

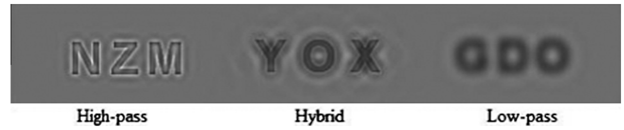

Fig. 1. Examples of filtered letters for the high-pass, low-pass, and hybrid conditions, respectively.

The previous studies examined either the impact of crowding on the spatial frequencies underlying identification performances (Chung & Tjan, 2007; Hess et al., 2000a, 2000b) or how target–flanker similarity, in terms of spatial frequency content, modulates crowding (Chung, Levi, & Legge, 2001). The aim of the present study is rather to examine how the different ranges of spatial frequencies contained in letters interact with letter confusability in a crowding paradigm. More specifically, targets and flankers were presented using one of the following spatial frequency filtering conditions: high-pass, low-pass, hybrid and broadband. In the case of the hybrid filter, the highest and lowest spatial frequencies remained whereas the middle, most useful, frequencies for letter identification were removed. Target–flanker distance was also manipulated, as well as target–flanker confusability.

2. Method

2.1. Observers

Twelve observers, aged between 19 and 23 (3 males and 9 females), with normal or corrected-to-normal vision, took part in the study. They all received monetary compensation for their participation and they were blind to the goals of the experiment.

2.2. Display

Stimuli were presented on a 17-in. DELL monitor with 1024 × 768 resolution at a distance of 57 cm from the observers. The experiment was controlled and programmed using MatLab (MathWorks, Natic, MA) with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Stimuli were uppercase 40 pt. Arial letters, subtending 1° of visual angle[2].

Using the Signal Processing Toolbox for MatLab, Butterworth filtering was applied to the stimuli to manipulate their spatial frequency content. The low-pass cut-off was at 1.61 cycles/letter and high-pass cut-off was at 3.14 cycles/letter. Crucially, these cut-off values were matched in terms of the capacity of the residual information (i.e. that remaining in the stimulus after filtering) to support the identification of single uppercase letters, based on the results of Fiset et al. (2008). The low-pass filter let through the low spatial frequencies of the stimulus but blocked those above the 1.61 cycles/letter cut-off. Conversely, the high-pass filter blocked spatial frequencies below the cut-off of 3.14 cycles/letter. The hybrid filter blocked the intermediate spatial frequencies (known to be the most important to support letter recognition) between the two cut-offs. Fig. 1 shows examples of spatially filtered stimuli. A broadband (non-degraded) version of each stimulus was also rendered. All conditions were matched in terms of stimulus

[1] The letter confusability scores were obtained by averaging the uppercase letter confusion matrices published in Van Der Heijden, Malhas, and Van Den Roovaart (1984), Loomis (1982), Gilmore, Hersh, Caramazza, and Griffin (1979), and Townsend (1971). They correspond to the total error rates for each individual letter of the alphabet. These values range between .24 (for the letter L) and .71 (for the letter B), with an average of .47 and a standard deviation of .13.

[2] Available confusability values are for uppercase letters only, thereby preventing the use of lowercase letters in the present experiment.