Source confusion is a major cause of crowding

Hans Strasburger

Department of Medical Psychology & Medical Sociology, University of Göttingen, Germany
GRP and Medical Psychology, University of München, Germany

Maka Malania

Institute of Cognitive Sciences, Agricultural University of Georgia, Tbilisi, Georgia
Laboratory of Vision, Beritashvili Center of Experimental Biomedicine, Tbilisi, Georgia

The loss of positional information for whole letters is one of the most important factors contributing to impaired letter and word recognition. Here we study the quantitative characteristics of flanker confusions in a crowding paradigm and test whether transient spatial attention relieves the crowding effect by reducing flanker confusions. We examined the crowding effect at three eccentricities for a range of flanker distances and attentional cue sizes. The effects of flanker distance confirm earlier findings that errors of both content and position are highest with flankers close by. However, the cue has no effect on flanker confusions and affects content information only, by enhancing target contrast sensitivity independent of cue size. Confusions with the inward, but not the outward, flanker increase linearly with eccentricity. Inward-flanker confusions dominate unlike reported asymmetries for masking. Our results are a psychophysical counterpart to separate neural coding of what and where in pattern recognition. The dependencies of cue effect and confusions on flanker distance scale with eccentricity and can be described by a generalized Bouma critical-separation rule. That rule shows a formal analogy to M scaling, from which the critical crowding distances on a cortical map can be derived as a logarithmic function. The perceptual results are visualized in a “doughnut” model.

Keywords: peripheral vision, visual attention, character recognition, contrast sensitivity, cortical magnification, perceptual organization, logarithmic mapping


Introduction

Origins of crowding theory

In 1936 the Danish ophthalmologist Ehlers first described what he later called “crowding” (Ehlers, 1936, p. 62; Ehlers, 1953, p. 432) and which was referred to as the “crowding effect” by Stuart and Burian (1962). Ehlers noted, in the context of normal reading and use of letter acuity charts, that there are visual, noncognitive difficulties of recognizing letters distant from fixation among other letters. He observes that the number of letters recognized is independent of angular letter size at varying viewing distance (Ehlers, 1936, p. 62). A natural way to study that effect is thus measuring the recognizability of a letter surrounded by others, and this is indeed the paradigm widely used nowadays (see Figure 1).

In a different line of reasoning, the German Gestalt psychologist Korte (1923) was the first to study the perceptual phenomena of reading in indirect vision. In his 66-page treatise, Korte meticulously describes reading errors and subjects’ reports of their perceptual experience, for letters and words, to arrive at a systematic description of the phenomena in pattern recognition when “seeing sidelong” (see Strasburger, Rentschler, & Jüttner, 2011, for a brief summary of his account). Out of the seven “causes of misreading” that he lists, five are perceptual phenomena related to impaired letter and word recognition. None of them perhaps describes crowding in today’s sense but they might be likened to processes and mechanisms underlying it. One of these, “word compression” or “perceptual shortening” of the perceived pattern (Korte’s Process e), has been widely recognized (Tyler & Likova, 2007, citing Pelli, Palomares, & Majaj, 2004, p. 1139; Levi, 2008); it describes the experience that part of the pattern within a word is lost. Here we wish to concentrate on another one of these: false localization of whole letters (part of Korte’s Process b). Like “perceptual shortening” it is described to occur in a dynamic phase of Gestalt formation, where “perceptions are . . . constantly moving, (with) bouncing of
whole characters ... (such that) subjects frequently speak of a 'dance'.” Korte also notes the special role of the letter borders: “Firm localization of detail is extremely difficult (and) is possible, at most, for the first and, less so, for the last letter.” In another of his examples, a particularly experienced subject in indirect vision saw, in the syllable “tot”: “a t and an o, but was unable to say whether the o was on the right or on the left, or whether there even was half an o on either side” (p. 41). We cite this example because it represents a vivid description of our data.

Current research on crowding

For many decades after Korte’s and Ehlers’ reports, crowding has been a niche research interest. Only more recently has it become more widely apparent, that it is one of the phenomena that distinguishes peripheral from foveal vision (for reviews see Strasburger, Harvey, & Rentschler, 1991; Pelli et al., 2004; Tyler & Likova, 2007; Levi, 2008; Strasburger et al., 2011). A number of mechanisms or causes have been proposed to underlie it. Levi (2008), e.g., distinguishes among others abnormal feature integration, loss of position information, (degeneration to) texture perception, configural grouping, and attentional influences. Here we will address three contributions to crowding: false whole-letter localization, altered contrast processing, and transient spatial attention. In particular, we are interested how these contributions interact and how they can be quantitatively described.

The contribution of whole-letter confusions in crowding (viz. lateral masking) has long been shown (Estes, Allmeyer, & Reder, 1976; Wolford & Shum, 1980; Strasburger et al., 1991; Huckauf & Heller, 2002; Chung, Legge, & Ortiz, 2003; Strasburger, 2005; Vul, Hanus, & Kanwisher, 2009). Estes et al. (1976), for example, showed that whole-symbol movement is far more likely to occur in the direction of the fovea. Wolford’s (1975) model of feature mislocalization included whole letters; Wolford and Shum (1980) differentiate between mislocalizations of features and errors involving whole symbols. Chung et al. (2003) presented a quantitative model for the dependency of letter source confusion on flanker distance. Our interest in that phenomenon was triggered by the results in Strasburger et al. (1991) and Strasburger (2005). In the latter we have shown that false localizations of letters (Korte’s “bouncing of letters”) contribute a large part to the crowding effect, and can be quantitatively separated from the neural processing of pattern content. Wolford and Shum (1980) similarly separate out the processing of pattern content when they speak of the failure to extract the required information from the display. The distinction between content and location can be seen as a psychophysical counterpart to the distinction of neural “where” and “what” processing in the dorsal and ventral streams (Ungerleider & Mishkin, 1982; Ungerleider & Haxby, 1994).

If the where is indeed processed separately from the content, the question arises how that locus is selected. Spatial attention could be involved here. Transient (bottom-up) and sustained (top-down) spatial attention are standardly distinguished (Nakayama & MacKeben, 1989); both have been proposed to contribute to crowding (Strasburger & Rentschler, 1995; S. He, Cavanagh, & Intriligator, 1996; C. He & Tjan, 2004; Strasburger, 2005; Fang & He, 2008). Spatial attention is known to heighten visual performance at the attended location. This has been shown with respect to many criterion measures (see Discussion for literature, and Strasburger et al., 2011, for review); Carrasco, Penpeci-Talgar, and Eckstein (2000) showed it for contrast thresholds, which are the criterion here. Strasburger (2005) therefore asked whether it is transient spatial attention—as evidenced by a transient positional cue presented at an optimum point of time before the target—that influences the coding of the where. Surprisingly, the transient cue left false localizations of letters fully unchanged, while it did reduce the crowding effect by improving the contrast thresholds for the target content. Transient attention thus interacts differently with those two contributors to the crowding effect, and is not related to the selection of the locus of information.

Figure 1. Stimulus arrangement. Cues are drawn to scale (for 4° eccentricity). On the left is the fixation point; the target (here a “7”) is shown as low-contrast white pattern on gray background at a certain eccentricity e in the right visual field. It is either shown singly or surrounded by a flanker left and right at distance d, measured from the center of the characters. The crowding effect is the increase of the target-recognition contrast threshold by the presence of the flankers. In the cued conditions, a circle of a certain size is shown shortly before the target.
Following previous work (Strasburger et al., 1991; Pelli et al., 2004; Strasburger, 2005), letter contrast thresholds are used here for measuring the crowding effect, i.e., the minimum contrast required for letter recognition. A few facts on that should thus be recalled. Letter contrast (recognition) thresholds are higher than grating contrast (detection/discrimination) thresholds (e.g., Solomon & Pelli, 1994; Petkov & Westenberg, 2003) and increase at a steeper rate with retinal eccentricity (Strasburger et al., 1991; Strasburger, Rentschler, & Harvey, 1994; Strasburger & Rentschler, 1996; Strasburger, 2001; Strasburger, 2003). Conceptually this might be accounted for by assuming “combination thresholds” (Strasburger et al., 1994) which the feature-encoding neurons (Fujita, Tanaka, Ito, & Cheng, 1992) need to exceed when feeding to the next stage. Note that such a framework lends itself to computational vision modeling (as requested by Tyler & Likova, 2007), where contrast information is transmitted along the visual pathway. Tjan and Nandy (2006) propose a technique for identifying features independent of their location which fits that framework and which the authors also applied to crowding (Nandy & Tjan, 2007). Unlike Wolford (1975), we therefore conceptualize feature integration as part of contrast processing. Crowding is then operationally defined as the increase of the letter contrast threshold by the presence of flankers. For simplicity we used flankers of the same contrast as the target, unlike Pelli et al. (2004) who also included the situation where the contrast of target and flankers was different. An advantage (over high-contrast letter crowding paradigms) of using contrast as the criterion is that other stimulus parameters (size, duration, locus, flanker distance, reaction time) can all be varied independently. Since it is an unspeeded task with only one choice, higher-level cognitive errors like erroneous response selection are less likely.

Wolford (1975) and many later models accounted for letter mislocalizations by feature perturbations, where features from nearby letters intrude into the target’s percept—a process termed source confusion. Wolford’s model used a feature set with seven types of features, including vertical lines, acute angles, and continuous curves. However, the data in Strasburger (2005) suggest that features are not floating arbitrarily and that those that constitute a letter often move jointly, i.e., are somehow bound together (as in Gestalt closure). Research on form vision also indicates the existence in visual cortex of cooperative mechanisms that locally connect feature analyzers (e.g., Shapley, Caelli, Grossberg, Morgan, & Rentschler, 1990; Lee, Mumford, Romero, & Lamme, 1998; see Strasburger et al., 2011, for review). In the context of pattern recognition we thus distinguished between mislocalized features and mislocalized whole letters, which we termed feature-source confusion and letter-source confusion, respectively (Strasburger et al., 2011).³

### Asymmetry

Crowding, flanker recognition, and flanker confusions are asymmetrical with respect to whether flankers are located towards or away from the fovea. An early observation comes from Wagner (1918, p. 52) who reports that, even though the first letter in a word is seen best, “interestingly at times in particular the last characters of a long word were identified.” Korte (1923, p. 41) writes that (in indirect vision) “stable localization of detail is extremely difficult (and) is possible at most for the first and, less so, for the last letter.” And later (p. 63) “that the interior of a word is identified as the last.” Mackworth, (1965, p. 68) reports that peripherally located flankers are more effective and recognition performance for a target letter in indirect vision doubles when a close-by masking flanker is moved from the peripherally to the centrally adjacent position. Bouma (1970, p. 178) notes from pilot data that the interference seems stronger from the peripheral side, and that “the area of interaction is . . . not quite circular around the position of the unknown letter but, rather, egg-shaped towards the retinal periphery.” Estes and Wolford (1971) reported that replacing a peripheral neighbor by an unchanging mask character (or by a blank space) had a greater facilitating effect on reportability of a target letter than replacing a central neighbor. Estes et al. (1976) showed flankers moved more often towards the fovea. Krumhansl (1977) found performance on the peripheral letter of a letter pair being higher than on the central letter, which implies more interference from the peripheral than from the central letter. Chastain (1983) found more interference when the flanker was peripheral, rather than central, to the target position (cf. also Bex, Dakin, & Simmers, 2003).

However, while the evidence that the peripheral flanker exerts more interference appears unequivocal, in certain other ways the central flanker seems to be the more important: Chastain (1982) reports that “confusability between members of a parafoveally exposed pair of letters affected accuracy of identifying the peripheral, but not the central, letter.” This implies that the similarity (to the target) of the flanker located centrally, not peripherally, plays a larger role for target identification. Chastain (1982) also points out that this result is incompatible with the explanation that interference results from feature movement from the peripheral to the central letter, since then similar flankers should exert less, not more, influence. A re-examination of Krumhansl’s (1977) data by Chastain (1982) further indicated that, similarly, performance on central targets in Krumhansl’s study was almost identical on confusable and nonconfusable trials, whereas performance on peripheral targets was much lower on confusable trials than on nonconfusable ones.
Thus, in these two studies the similarity of the flankers to the target—i.e., the influence of flanker features on the target—showed the opposite asymmetry than that found for interference. Huckauf and Heller’s (2002) results confirm that, for confusions between target and flankers, the inward flanker plays the more important role. Huckauf and Heller further found, however, that a strong left-right effect (i.e., of reading direction) is superimposed on that so that the resulting asymmetry of the rate of confusion is high in the right hemifield but weak in the left field. These studies further imply that floating of features is not the only mechanism underlying crowding.

More recently, Petrov and Popple (2007) remark, for orientation crowding, that the outward Gabor can be most easily identified by all observers. van den Berg, Roerdink, and Cornelissen (2007, suppl. figure 7) could not replicate that result but instead found a robust asymmetry for letter crowding. Petrov, Popple, and McKee (2007) report that an outward mask (of same contrast as the target) is, on average, four times more disruptive than an inward mask for orientation identification, but not for detection, at low target contrast with high-contrast flankers. From their findings they ascribe particular significance to the inward–outward anisotropy and suggest it could be the hallmark of crowding. Petrov and Meleshkevich (2011) demonstrate that this stronger masking by the outward flanker can be ascribed to the influence of spatial attention which is commonly mislocalized outward of the target, and that it disappears when attention is diffused. We were thus interested whether the data on flanker confusions, in Strasburger (2005) and those presented here would support an anisotropy, and asked in which direction a potential asymmetry would go with respect to confusions.

Goal of the paper

The results in Strasburger (2005) left open the question whether the unexpected results of the study were specific to the chosen ring cue (which had a size that was M-scaled with eccentricity). There was further a slight inconsistency in the data that could have been connected to the cue size. So in the present paper we varied, additionally to eccentricity and flanker distance, the cue size in a large range and also considerably extended the data base. The aim of the study was to examine (a) whether an increasingly larger cue has a changed effect on contrast processing, (b) whether at some other size the cue does, after all, have an effect on the positional errors (the “where”), and (c) by what analytic function on flanker distance the extent of the effects on contrast threshold and on flanker confusion can be described. It turns out to be an inverse-U shaped function for both contrast thresholds and flanker confusions. Beyond the peak, the function for the contrast thresholds is log linearly dependent on flanker distance, and for confusions it is linearly dependent. The functions are largely described by their respective location of the maximum and the vanishing point on the abscissa. So, the final question was whether and how these characteristic points scale with eccentricity. The answer is stated as a quantitative, generalized Bouma rule. We use Bouma’s original definition of flanker distance in this since only then are the equations well-defined in the fovea (Strasburger, 2005). To appreciate the description’s consequences, the function is visualized in 2D visual space as a “doughnut” of inhibition versus enhancement.

Methods

Contrast thresholds for the recognition of characters were obtained in a 10-afc adaptive procedure based on the ML-PEST algorithm (Harvey, 1986, 1997). The procedure runs on PC hardware (Irtel, 1997; Jüttner & Strasburger, 1997) and uses dithering to achieve 10-bit gray scale resolution (Bach, Meigen, & Strasburger, 1997). Thresholds are defined as the point of inflection on a Weibull function (67% correct in a 10-afc task; note that by this the error rate at threshold is constant at 33%). The target stimulus was one of the 10 digits (0–9), presented for 100 ms as white pattern on a gray background of 50 cd/m² luminance (Figure 1). Pilot experiments with letters instead of numerals had given similar results so we stayed with the digit stimulus set for compatibility with previous findings. Additional flankers from the same set and ring cues were presented as described below. Digits were rasterized in multiples of a 5×7 pixel matrix and their size was specified as character height; so the width w of a character is 0.71 times the specified size. Stimulus contrast was varied from trial to trial (in log Michelson units) as specified by ML-PEST, with all other conditions unchanged. A run was ended when the estimated threshold confidence interval reached a prespecified level of 0.2 log units (i.e., ±0.1). On average this took about 25 trials (25.0 ± 0.31 SEM, estimated from two subjects with 11,700 responses). For a sample run see Strasburger (2001, figure 3). Contrast is specified as percent Michelson contrast or the log thereof throughout. Stimuli were displayed on a 19-inch Iiyama monitor (Iiyama Deutschland GmbH, Rohrbach/Ilm) (1024 × 768 pixels), calibrated in size and luminance by the PXL software (Irtel, 1997) using a Mavo-Monitor digital luminance meter (Gossen GmbH, Nürnberg). The monitor was framed by a large white cardboard panel, and room lighting was set such that screen and frame.
had similar luminance to have stable photopic adaptation. Ten young observers of both genders (age 18–36) with normal or corrected-to-normal visual acuity (FrACT ≥ 1.0, Bach, 1996) were tested in identical conditions; each condition was measured twice. For in-depth discussion of the methods see also Strasburger et al., 1991; Bach et al., 1997; Strasburger, 2001; Strasburger, 2005). Note that, unlike in most other paradigms used for studying crowding, error rate in the contrast-threshold paradigm is essentially fixed to one minus percent-correct at threshold, 33% in this study. Error rate is thus not a criterion for crowding here.

Three basic conditions were tested: (a) in a single-character condition (i.e., unflanked and uncued) serving as baseline, only the target was presented at a certain eccentricity on the right meridian. (b) In a flanked, uncued condition the target was surrounded by a neighboring character left and right of same contrast, at one of a set of flanker distances (note that this was different from Pelli et al.’s, 2004 paradigm where flanker contrast is held constant); (c) in the cued conditions there were flankers as before, and additionally a circle appeared 150 ms before the stimulus triplet to attract transient attention to its center (Eriksen & Collins, 1969; Nakayama & MacKeben, 1989). Since target and flanker locations were fixed by design there was no uncertainty in the paradigm as to where the circle and target would appear. Stimulus sizes were M-scaled with $E_2 = 2.0^\circ$ (Levi & Klein, 1985); cue sizes were chosen in approximately log scaled steps to span a wide range. Target eccentricity, flanker distance, and cue diameter were all varied independently. Subjects fixated a dark fixation point displayed on the screen continuously during the stimulus presentation. Table 1 summarizes stimulus conditions.

The target measures were contrast threshold and percentage of correspondences between a subject response and a flanker (%-correspondences), the latter defined as those errors where the observer reports a flanker instead of the target. Chance level of correspondences is $p_c = 0.1 + 0.1 - (0.1 \times 0.1) = 19\%$ (since left and right flanker each have a chance or 10% to be reported and in 1% of cases are equal by chance; the chance rate was verified by a Monte Carlo simulation). A total of close to 55,000 trial responses were obtained.

### Results

#### Contrast thresholds

Figure 2 shows the effect of cue size on the crowding effect at the three eccentricities. Depicted are the target threshold contrasts as a function of flanker distance (top blue line) in comparison to single performance (dashed horizontal line); the lines in between show cued performance for the set of cue sizes. As expected, the crowding effect increases with decreasing flanker distance, and the cue does have an impact on contrast threshold. Importantly, however, there is no noteworthy effect of cue size. The small variations between the lines for the large range of cue sizes can be attributed to chance. This was confirmed by three one-way ANOVAs with cue size as factor, which resulted in $p$ values between 0.75 and 0.97.

We therefore averaged the log cued thresholds across cue sizes (Figure 3), and calculated the difference between (log) uncued and cued thresholds. This gives us a measure of mean cue effect, or contrast gain [since $\ln(c_1) - \ln(c_2) = \ln(c_1/c_2)$]. Note that this is not a pure measure of the cue’s effect in general; it is constrained by the crowding-induced loss of contrast sensitivity and reflects the cue effect in a letter crowding situation. It is shown in Figure 4 as a function of flanker distance. The decrease with flanker distance would be expected (since crowding decreases) but the regularity of the rate at which this happens is striking. This allows reliable estimation of the point at which the observed cue effect vanishes (by extending the regression line to the right), i.e., the critical flanker distance. Note further that, even though cue gain is constrained by crowding, these critical values are not necessarily smaller than crowding critical distances since they are based on the range where the cue effect is nonvanishing. Comparing them to Figure 3 it turns out they are indeed a little larger; the predicted values are 1.7°, 3.1°, and 4.5° at the three target eccentricities. Expressed as a fraction of the latter they are 85%, 78%, and 75% of eccentricity.
respectively, with a mean of 79% of eccentricity. Bouma (1970) showed that the critical flanker distance in the crowding effect for high-contrast letter recognition scales with eccentricity, and, as a rule-of-thumb, amounts to approximately 50% of target eccentricity. With 79% rounded-up to 80%, the presented results could thus loosely be stated as an “80%-Bouma rule” for the effect of a transient cue in crowding. We will discuss more precise statements on Bouma’s Law below.

A further interesting characteristic of the observed cue effect’s relationship with flanker distance seen in Figure 4 is that it is not ever increasing with flanker distance, with a mean of 79% of eccentricity. Bouma (1970) showed that the critical flanker distance in the crowding effect for high-contrast letter recognition scales with eccentricity, and, as a rule-of-thumb, amounts to approximately 50% of target eccentricity. With 79% rounded-up to 80%, the presented results could thus loosely be stated as an “80%-Bouma rule” for the effect of a transient cue in crowding. We will discuss more precise statements on Bouma’s Law below.

Figure 2. Effect of cue size (traces) on contrast threshold versus flanker distance, at 2°, 4°, and 6° target eccentricity (from top to bottom). Top blue line: target contrast thresholds in the flanked, uncued condition; dashed horizontal line: unflanked, uncued performance (labeled as single); cued performances are the lines in between. The crowding effect is seen as the vertical distance between the top line and dashed line; the cueing effect is the reduction of this distance. Note the log scaling of the contrast axis.

Figure 3. Cued (open symbols), and mean uncued (filled symbols) contrast thresholds versus flanker distance for the three eccentricities tested (blue traces: 2°, red traces: 4°, black traces: 6°). The three single-digit thresholds (dashed blue horizontal line with error bars) are very similar.

Figure 4. Mean cue effect on contrast threshold (viz. contrast threshold gain) versus flanker distance, for the three measured eccentricities (traces).
distance decreasing, but reaches a maximum below which the effect steeply decreases. The three maxima are at 0.5°, 0.75°, and 1°, respectively, i.e., again they scale with eccentricity. Expressed as a fraction of the latter, the values are 25%, 19%, and 17% of eccentricity or, on average, 20%. So, not only the upper limit of interaction (the critical distance) but also the lower limit scales with eccentricity. We might call it a “Bouma range” of the observed transient-cue effect in a crowding task. The range and the function’s shape scale with eccentricity, setting in at a little below 20% to 80% eccentricity and having a maximum at 20% eccentricity from where it decreases log-linearly to the vanishing point at 80% eccentricity. The contrast gain factor in that maximum ranges between 1.5 and 1.7, i.e., there is a substantial increase of the target’s effective contrast induced by the transient cue.

Localization errors

The next part of the analysis concentrates on localization errors, i.e., on the number of cases (out of the constant error rate) when the subject’s response corresponds not, as instructed, to the target but to one of the two flankers surrounding the target. Figure 5 shows how these correspondences depend on flanker distance for the set of cue sizes at the three target eccentricities (2°, 4°, and 6°) (left and right flankers are lumped together here; they will be analyzed separately, below). Strikingly, correspondences depend very systematically on flanker distance, whereas the presence of a cue has no systematic effect whatsoever, independent of its size. This was again confirmed by three one-way ANOVAs with cue size as factor, which resulted in \( p \) values between 0.46 and 0.79. So, whereas the cue did influence the target contrast threshold, as shown in the previous section, it did not affect letter source confusion.

Since the variations between the uncued and the various cued performances seem random, it is (again) meaningful to take their averages. These averages of %-correspondences over uncued and all cued conditions, versus flanker distance, are shown as bold black curves in Figure 5. They are redrawn in Figure 6, together with regression lines for the high flanker-distance limb. Note that the correspondences unlike contrast thresholds are drawn in linear scale. Five conclusions (a–e) may be drawn from Figure 6: (a) Letter source confusion decreases with flanker distance from a certain point onward, which is as expected; (b) letter source confusion increases with eccentricity, and with increasing eccentricity occurs at larger flanker distances, i.e., it scales with eccentricity; (c) again, a critical distance can be reliably determined by linear regression.
For the three eccentricities the critical distance values are 1.7\(^8\), 3.05\(^8\), and 4.2\(^8\), respectively. Expressed as a fraction of eccentricity these correspond to 85\%, 76\%, and 70\%, respectively, i.e., they again scale with eccentricity and can be summarized as about 81\% \times eccentricity on average (note that we will recalculate these values differently in the Discussion section under Bouma’s Law). This value is very similar to the critical distance found for the cue effect on contrast thresholds, which we will also take up in the Discussion. (d) As with the cue effect on contrast thresholds, there seems to be a maximum of the cue effect on letter source confusion, occurring at 0.45\(^8\), 0.8\(^8\), and 1.2\(^8\), respectively. Expressed as fraction of eccentricity the maxima are at 22\%, 20\%, and 20\% of eccentricity (on average 21\%). (e) At the maximum there are up to 38\% false localizations (with error rate constant by design). Subtracting the chance level of 19\% leaves 19\% false localizations above chance. So every fifth percept at this close flanker distance has lost its correct position code.

Inward/outward asymmetry

The large amount of data that we have acquired on report/flanker correspondences allows us to shed light on a longstanding related question—what underlies the perceptual asymmetries regarding the role of inward vs. outward flankers? Would such crowding asymmetries show themselves in the position confusion behavior? In our early data (Strasburger et al., 1991, Table 2) we had not found asymmetry in positional errors at 4\(^8\) eccentricity but now, with more data at hand and following a suggestion of J. Solomon, we were interested to look for asymmetries in our new data. We thus reanalyzed both our Journal of Vision 2005 data (which constitute a large dataset of 34,000 responses; Strasburger, 2005) and the present data. Results are shown below in Figure 7 and 8, respectively. Since, in these two studies, the crowding task was performed on the right meridian only, we cannot distinguish between an inward/outward or left/right asymmetry from these data, so the asymmetries.

Table 2. Foveal critical spacings and \(E_2\) values for critical spacing, from Equations 5–8 plotted in Figure 9.

<table>
<thead>
<tr>
<th>Condition</th>
<th>(d_0)</th>
<th>(E_2)</th>
<th>Eq.</th>
<th>Figure 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical distance, thresholds</td>
<td>0.3(^8)</td>
<td>0.43(^8)</td>
<td>5</td>
<td>Pink</td>
</tr>
<tr>
<td>Maximum, thresholds</td>
<td>0.25(^8)</td>
<td>2.0(^8)</td>
<td>6</td>
<td>Blue</td>
</tr>
<tr>
<td>Critical distance, correspondences</td>
<td>0.48(^8)</td>
<td>0.77(^8)</td>
<td>7</td>
<td>Pink</td>
</tr>
<tr>
<td>Maximum, correspondences</td>
<td>0.07(^8)</td>
<td>0.37(^8)</td>
<td>8</td>
<td>Blue</td>
</tr>
</tbody>
</table>

Figure 6. Mean correspondence data (% correspondences of wrong response with either flanker) versus flanker distance, fitted with a regression line on the right limb. Data are redrawn from the bold curves in Figure 5 A–C, for the three target eccentricities.

Figure 7. Correspondences of the subject’s report with the left or right flanker in the data of Strasburger (2005); (a) shows the correspondences separately for left and right, each for the three conditions flanked, cued, and content-only (see text); (b) shows the resulting left-right ratios of flanker correspondences. Note that the eccentricity range was different in that study.
reported below could be either (and in light of Huckauf & Heller, 2002, are probably a combination of the two). These analyses yielded similar results for the two studies.

In the Journal of Vision 2005 data (Strasburger, 2005), three conditions were compared: a flanked and a cued condition were identical to the present conditions; the third was a flanked, uncued condition, labeled content-only in the paper, where the thresholding algorithm was changed slightly, arriving at slightly lower contrast thresholds (the latter data are shown for completeness only). It turned out that a left-right asymmetry depended on the target’s (right horizontal) eccentricity (Figure 7): There is a systematic increase of correspondences for both left and right flankers and for all three conditions (Figure 7a). That increase, however, is slightly steeper for the left characters so that the left-right ratio increases above one (Figure 7b). Inward (or left) characters were thus more easily mistaken for the target with increasing eccentricity. Paradoxically, this is most pronounced for the cued condition, where the target location is most clearly marked.

In the new study cue size was varied along with eccentricity and flanker distance. That allowed analyzing how the changes in cues size influence the left-right ratio of flanker correspondences. Since the ratio changed little with cue size, however, (with a slightly larger ratio at intermediate cue sizes for 2º and 6º eccentricity), we collapsed the data over that parameter. The results are shown in Figure 8. The effect turns out similar to that observed in the previous study (Strasburger, 2005) but was even more pronounced: The correspondences with the outward (the right) flanker are fully independent of eccentricity whereas those with the inward (the left) flanker increase linearly and reliably with eccentricity. The constant terms are close to chance level of correspondences (19%) as would be expected:

\[
c_{\text{inw}}(\%) = 2.97E + 15.8\% \\
c_{\text{outw}}(\%) = 19.8\% 
\]

As a result, the overall left-right ratio shows a robust increase with eccentricity (Figure 8B), where the inward flanker (the left flanker) but not the outward flanker is increasingly often confused with the target (Figure 8A).

### Discussion

#### Summary of results

Here are the results of the present study that we will discuss in the following:

1) Contrast thresholds and positional errors both decreased with flanker distance (Figures 2 and 3, and Figure 5);
2) The cue had a gain-control effect on contrast thresholds (Figure 4), and had no effect on positional errors;
3) The gain control effect (as observed, i.e., constrained by the crowding-induced loss of contrast sensitivity) scaled with eccentricity (Figure 4);
4) The cue’s effect was independent of cue size (Figure 5);
5) The functional relationship between correspondences and flanker distance scaled with eccentricity (Figure 6), i.e., was shifted towards larger flanker distances when eccentricity increased;
6) Bouma’s rule can be extended to describe the point of maximum cue gain-control effect (as observed, i.e., constrained by the crowding-induced loss of...
contrast sensitivity) and the critical distance thereof (Equations 5 and 6);
7) Bouma’s rule can be extended to describe where the maximum of positional errors occur and the critical distance of these (Equations 7 and 8);
8) Bouma’s rule, when stated with non-zero intercept, has a formal analogy with cortical magnification (Equations 10 and 11; a prerequisite for the next);
9) Bouma’s rule can be mapped onto the primary visual cortex (Equations 13 to 15); that analysis is required to show that the flanker confusion asymmetry is compatible with the cortical map;
10) These characteristics can be visualized and memorized with a “Doughnut model” (Figure 10);
11) Confusions of target and flanker were asymmetric, with the inward (leftward) flanker playing the more important role, opposite to the asymmetry found for flanker interference (implying that separate mechanisms underlie). Confusions with the inward (but not the outward) flanker increased with eccentricity.

The Discussion is organized to first deal with the two mechanisms that are covered (spatial attention and letter source confusion), ask how independent the two measures are that were considered, next discuss questions regarding asymmetry, then revisit Bouma’s rule and develop a quantitative description of our data by it. To ask which direction of asymmetry is compatible with cortical architecture we derive equations for the cortical retinotopic map of Bouma’s rule. Finally we summarize our data in a descriptive model.

Spatial attention

Spatial attention has gained wide interest in the context of crowding. Spatial covert attention (i.e., without eye movements) –which is at issue here—is an aspect of visual selective attention (for reviews see, e.g., Bundesen, 1990, 1998; van der Heijden, 1992; Schneider, 1993; Desimone & Duncan, 1995; Gazzagna, 1995; LaBerge, 1995; Yeshurun & Carrasco, 1999; Chalupa & Werner, 2004; Strasburger, 2005; Carrasco, 2011; for computational models of overt attention and saliency see Chalupa & Werner, 2004, and Bruce & Tsotsos, 2009). For letter crowding, Wolford and Chambers (1983) were, to our knowledge, the first to...
The difference is that the attention-attracting cue appears before the target and enhances performance in its area, whereas the maskers for the flankers appear after the target and suppress performance in their area. An interesting aspect of visual spatial attention has been brought up by Tsotsos (1990) (followed up, e.g., by Mounts, 2000; McCarley & Mounts, 2007; McCarley & Mounts, 2008 and modeled by Tsotsos et al., 1995, and Cutzu & Tsotsos, 2003; see Carrasco, 2011 for review): Whereas the “spotlight of attention” is typically assumed to decay monotonically around its center, there may be an inhibitory surround. Mounts’ results, following Tsotsos’ prediction, show an inhibitory annulus that is of limited extent and shows an inversion further out, reminiscent of a “Mexican hat.” Our (present and previous) data only show decay in the effect, as do crowding data of other studies. From these data it thus does not seem that the inversion seen for spatial attention finds a parallel in the context of crowding, but it will be interesting to pursue this possibility further.

### Letter source confusion

With respect to the mechanism of letter-source confusion we need to address the following question: “What is the difference between individual features floating and a whole letter floating?” In a featural approach like that of Wolford (1975), the perception of a letter arises if a majority of the detected features are characteristic of that letter. So if most of the constituting features float, the entire letter floats. The difference between the two confusion mechanisms is in the way features are bound. In feature-source confusion, individual features lose their position code, i.e., they lose their marking as to which character they belong to. This is what is addressed in Korte’s first part of phenomenon (b), in Wolford (1975), Krumhansl and Thomas (1977), Saarinen (1987), Pelli et al. (2004, p. 1137), and in Tyler & Likova, 2007 Figure 2a). In the second case—letter source confusion—the features keep their marking as to which character they belong to, remaining bound to each other, but the entire character loses its position code. This is the phenomenon Korte (1923) was originally describing to whom he spoke of the “dance” (Korte’s second part of b). Grouping accounts, like those of Livne and Sagi (2007; 2010) or May and Hess’ (2007) “snakes & ladders,” fit into that framework as providing the glue by which features are bound. The Gestalt concept of closure refers to the same phenomenon. Chung and Legge (2009) recently presented a quantitative model for the extent of letter position confusions with varying eccentricity. Zhang et al. who worked with Chinese characters speak of letter-level position uncertainty; they show how positional errors can be separated from

The attentional crowding-relieving effects discussed here need to be conceptually distinguished from those in the study of Chakravarthi and Cavanagh (2009), who showed relieving of crowding by masking the flankers. The attentional crowding-relieving effects discussed here need to be conceptually distinguished from those in the study of Chakravarthi and Cavanagh (2009), who showed relieving of crowding by masking the flankers.

Spatial attention (also called covert attention) is often operationally defined as the influence of a spatial cue, as described by Eriksen in the 1970s (Eriksen & Collins, 1969; Eriksen & Rohrbaugh, 1970; Eriksen & Hoffman, 1974). It is commonly divided into two kinds, sustained and transient (Nakayama & MacKeben, 1989), or, similarly, into voluntary and automatic attention (Jonides, 1981; Yantis & Jonides, 1984). In previous work (Strasburger & Rentschler, 1995; Strasburger, 2005) we surrounded the target with a ring cue which was presented at an optimum time before target onset (Eriksen & Collins, 1969) to attract transient attention. This produced an interesting differential effect: while the cue did substantially improve recognition performance, it left positional errors fully unchanged. This improvement in recognition is evidence that transient spatial attention is concentrated at the target, either by enhancing neural activity at the target position or by suppressing activity at neighboring positions. At the same time, it is evident that transient attention does not provide the sought-after selection mechanism for the letter source.

The present results now confirm and extend the findings of our previous study. We demonstrated that contrast processing is enhanced by the transient ring cue without influencing letter source confusion, and both these effects are independent of cue size (Figures 4 and 5). The ring cue’s role is that of enhancing the coding of content by increased transient attention (cf. Carrasco et al., 2000)—i.e., providing a “brighter spotlight”—while leaving position coding of the flanking characters unaffected. Yeshurun & Carrasco (2008), who find little effect of cue size in a texture segmentation task, speak of the rigid, fast and automatic nature of transient attention that does not adapt its operation on the basis of the size of the attentional cue. The (imperfect) selection of the target over the flankers is governed by sustained (voluntary) attention, since subjects are aware in advance of where the stimulus will appear.

The attentional crowding-relieving effects discussed here need to be conceptually distinguished from those in the study of Chakravarthi and Cavanagh (2009), who showed relieving of crowding by masking the flankers.
identity errors (Zhang, Zhang, Xue, Liu, & Yu, 2009) and, by comparing partial and whole-report paradigms, further break down whole-letter localization errors showing position swapping between target and flankers (Zhang, Zhang, Liu, & Yu, 2012).

Do positional errors and the cue effect just mirror the crowding effect itself?

The shape by which positional errors and contrast gain vary with flanker distance look rather similar and the question arises how they are related. Do they just mirror crowding or are they independent effects? Short of a final answer based on the present data, a few observations are in place. Cue-induced contrast gain is clearly constrained by crowding as defined here, i.e., defined as a reduction of target contrast sensitivity by the presence of flankers. When there is no such reduction (since flankers are beyond the critical distance) it cannot be further reduced. By this token, we would see in the data if the area within which the underlying mechanism acts were smaller (which is not the case) but not if it is larger (the critical distances that we found for the gain were in fact a little larger than those for crowding since they resulted from extrapolation). However, whether the shape of the decline within the critical region is that of the underlying mechanism (is a “true” gain improvement) or that of the crowding effect cannot be answered based on the present paradigm. The metric that we used for assessing the cue effect was

\[
\log(C_{\text{cued}} / C_{\text{uncued}}) = \log(C_{\text{cued}} / C_{\text{uncued}}),
\]

and for assessing the crowding effect

\[
\log(C_{\text{uncued}} / C_{\text{single}}) = \log(C_{\text{uncued}} / C_{\text{single}}).\]

Neither of these are ratio scales which would allow some kind of normalizing, and we do not believe a metric for the underlying effect could be derived in the present framework. So we stay with the overtly observable effect (Figure 4), in line with our descriptive goal.

As to the other observed variable—the correspondences or false localizations, which are defined as a specific kind of error—it should be pointed out that, unlike in typical other crowding paradigms, they are not constrained by the error rate. Here they rather represent a variable proportion out of an essentially unchanged error rate, that proportion increasing and decreasing with flanker distance. Since this kind of error is considered a constituent of the crowding effect we would not say it is constrained by it, but rather that this constituent of crowding shows, on a linear scale, a similar behavior than contrast-threshold gain on a log scale. For a theoretical link between the two the log transducer model comes to mind (e.g., May & Solomon, 2012) but this is beyond our scope. As to a common ground for contrast thresholds and false localizations one should also remember how the cue acts radically different on the two contributors to crowding.

Evidence in the literature speaks for a certain independence of whole-letter confusion from other sources in the crowding effect. A recent modeling study by Hanus and Vul (2013) emphasizes that contributions of both letter confusion and letter substitution are required for successful modeling of crowding. Zhang et al. (2009) show with Chinese characters (which are more complex than roman characters) that there is a difference between within-character and between-character crowding show how positional errors can be separated from identity errors. Zhang et al. (2012) speak of letter-level position uncertainty and separate-out different kinds of whole-letter localization errors.

Asymmetry

The data that we acquired on letter source confusion (Strasburger et al., 1991; Strasburger, 2005; Strasburger & Malania, 2008) allowed an analysis with respect to an asymmetry in the influence of flanking characters. We found confusions with the leftward / inward (but not the outward) flanker to be increasing with eccentricity, which led to increasing asymmetry with eccentricity (Figures 7 and 8). This spatial asymmetry of flanker confusion would correspond nicely to Petrov, Popple, and McKee’s (2007) characterization of crowding relative to surround suppression—if it were not for the fact that the asymmetry in our data is in the opposite direction to that commonly reported: the inward, not the outward, flanker was more often confused with the target, and these errors have the effect of decreasing target recognition, i.e., contribute to crowding. However, unexpected as it had been, our result strikingly confirms that there are cases of an opposite direction of asymmetry like in the long-neglected findings of Chastain (1982) stated above: Opposite to the standardly cited effects of flanker interference, with respect to the effects of similarity and confusion the inward flanker plays the important role, that importance increasing with eccentricity. Note that this conclusion implies that floating of features cannot alone explain the crowding effect since its asymmetry is opposite to that of confusions and confusability (Chastain, 1982). Note also that this complements rather than contradicts the results on interference (like those of Petrov et al., 2007, or Bouma, 1970): it just shows that the different contributions to the crowding effect can have different characteristics. It would also explain why results on the asymmetry are sometimes not clear-cut.
Bouma's law

Bouma (1970) has formulated an intriguing rule relating crowding to the distance of the flankers from the target. We will use and extend this rule for a descriptive model for the present results. Bouma’s rule states that the critical flanker distance \( d \) below which crowding sets in, when expressed as free space between the letters, is about 50% of the target’s eccentricity. Bouma’s rule thus states that

\[
d_c = bE
\]

where \( d_c \) is free space and \( b \) is 0.5. Nowadays, however, flanker spacing is commonly measured not as free space but as the center-to-center distance. Bouma’s rule then translates to

\[
d = bE + w
\]

where \( d \) is the center-to-center flanker distance, \( b \) is around 0.5, \( E \) is eccentricity in deg, and \( w \) is letter width in deg. Interestingly that relationship is not proportionality, as is commonly quoted; it is linear with a positive \( y \)-axis intercept. The intercept is equal to letter size. Though it seems a small difference, the presence of a nonzero intercept is important for consistency: Proportionality would be ill-behaved in the fovea since flankers would then superimpose with the target before they crowd. Bouma’s equation, in contrast, is well-behaved at small eccentricities. The slope is Bouma’s factor (around 50%) and the \( y \)-intercept is a prediction of the foveal critical crowding distance (when measured center-to-center).

Bouma’s rule was meant as such, as a rule of thumb to provide an overall picture of the spatial relationships. Indeed, empirical values for the slope constant in Bouma’s rule vary quite a bit as is documented by Pelli et al. (2004, table 4). One reason for that is that critical spacing is often not well defined, as seen, e.g., in Figures 2 or 3. Yet with Equation 4 the regular results in Figure 4 and Figure 6 allow highly reliable estimates through linear regression to the vanishing point (i.e., to the intersection of the regression line with the \( x \)-axis). The point is determined on a log-linear scale for the contrast thresholds and on a linear scale for the correspondences. The estimates from the Results section are summarized in Figure 9 and Equations 5–8:

\[
d_{\text{crit,thr}} = 0.7E + 0.3^\circ
\]

\[
d_{\text{max,thr}} = 0.125E + 0.25^\circ
\]

\[
d_{\text{crit,corr}} = 0.625E + 0.48^\circ
\]

\[
d_{\text{max,corr}} = 0.188E + 0.07^\circ
\]

The estimates show remarkably systematic eccentricity scaling. The steep top lines in Figure 9 are the critical distances \( d_{\text{crit}} \) (Equations 5 and 7), the shallow bottom lines the positions \( d_{\text{max}} \) of the effect maxima (Equations 6 and 8), both obtained from Figure 4 and Figure 6. Bouma values (in the original meaning as the slope \( b \) in Equation 4) turn out as 70% for the cue’s effect on contrast thresholds, and 63% for its effect on correspondences, i.e., are slightly smaller than the values under the proportionality assumption derived in the Results section. The slope values for the position of the effect maxima (blue line) are 12.5% and 19%, respectively. \( Y \)-intercepts are all positive and in the order of 20 minutes of visual angle (see below). Note that Equation 4 with its nonzero constant term fits the data much better than would proportionality (as in Equation 3).

Interestingly, Petrov and McKee (2006) found very similar relationships for Gabor-grating surround suppression (see above). The log-contrast threshold elevation functions were highly linear with target-surround separation for a range of conditions (ib. Figure 7)—reminiscent of Figure 3—so that interaction ranges could be reliably determined. Furthermore, Petrov and McKee’s interaction radii nicely obeyed a linear relationship with eccentricity, with a positive \( y \)-intercept of 0.41° (ib. Figure 8). From that figure the relationship is

\[
r = 0.1E + 0.41^\circ
\]

It is thus quantitatively comparable with the scaling behavior of the maximum cue effect on log-contrast thresholds described in Equation 6.

Formally, Equation 4 is equal to \( M \)-scaling; (cf. Levi & Klein, 1985; Strasburger, 2003, chapter 4.1; Strasburger et al. 2011; Drasdo, 1991). Using the notation introduced by Levi and Klein (1985), we can thus rewrite Equation 4 as

\[
d = d_0 \left(1 + \frac{E}{E_2}\right),
\]

where \( d_0 \) is the foveal critical spacing, and \( E_2 \) is Levi’s value where \( d_0 \) doubles. Equations 5–8 can then be restated according to this equation; the corresponding values for \( d_0 \) and \( E_2 \) are summarized in Table 2. Note that neither the \( E_2 \) nor \( d_0 \) values are reliable estimates since there are no measurements below 2° eccentricity, and slight deviations from linearity (e.g., Gurnsey, Roddy, & Chanab, 2011) have a large effect (cf. review in Strasburger et al., 2011). Foveal critical spacings are probably much smaller (e.g., Toet & Levi, 1992). These values are meant here as a rough shorthand description of the overall scaling functions.

Bouma’s rule mapped onto the cortex

Crowding is a cortical phenomenon (as first pointed out by Flom, Weymouth, & Kahnemann, 1963; cf.
Strasburger et al., 2011) and it is thus fitting to ask whether our finding of a reversed asymmetry fits with what we know about the retinotopic map in early visual areas (we thank an anonymous reviewer for pointing that out). To answer that question we need to derive Bouma’s rule in cortical terms (Equations 13–15). The reader not interested in cortical maps can skip these derivations.

M-scaling was originally conceived as reflecting the scaling properties of the primary, retinocortical visual pathway, and the inverse-linear cortical magnification rule can be restated using $E_2$ as

$$M^{-1} = M_0^{-1} \left(1 + \frac{E}{E_2}\right), \quad (11)$$

where $M_0$ is the foveal cortical magnification factor and $E_2$ is Levi’s value (at which $M^{-1}$ doubles in this case). Two implications of the M scaling analogy and retinotopic mapping should be considered: (a) given retinotopic mapping with an M scale: what is the locus of the flankers in the primary visual cortex? (b) Is an asymmetry of the flanker interaction compatible with this framework? Pelli (2008, Equation 3) derived an answer to question (a) which states that, if spacing in a crowding task is proportional to eccentricity (i.e., if it obeys Bouma’s rule as in Equation 3), then the flankers appear at fixed spacing at the primary visual cortex. Now, however, Pelli’s derivation is based on the logarithmic mapping (of visual space onto the cortex, $\delta \propto \ln E$) derived by Schwartz (1980), which Schwartz, for simplicity, based on the proportionality assumption. The logarithmic mapping has been repeatedly verified to hold for eccentricities larger than about 3° eccentricity but is not meant for and is not valid at small eccentricities (Strasburger et al., 2011; see there for the derivations of the following equations). It is, in particular, undefined in the fovea. Fortunately, the reasoning can be generalized by using the standard inverse linear cortical magnification rule (Equation 10). Under this rule, one can derive that (in analogy to Schwartz’s mapping) the cortical distance $\delta$ of the target from the fovea is

$$\delta = M_0E_2 \ln \left(1 + \frac{E}{E_2}\right), \quad (12)$$

with notations as before (note that, unlike the original logarithmic mapping, this equation is well-defined in the fovea). With Equations 10–12 one can then derive how critical spacing on the cortex, $\Delta \delta$ varies with eccentricity:

$$\Delta \delta = M_0E_2 \ln \left(1 + \frac{1}{d_0} \frac{E}{E_2}\right), \quad (13)$$

where $E_2$ refers to local size (Equation 11) or to distance from the fovea (Equation 12), and $E_2$ refers to critical spacing in Equation 10. Different symbols are used for $E_2$ from Equations 10 and 11 since the variation of local size and location on the one hand, and of critical spacing on the other hand are likely different. The behavior of Equation 13 quickly converges to the value

$$\lim _{E \to \infty} \Delta \delta = d_0 \left(1 + \frac{E}{E_2}\right), \quad (15)$$

e.g., becomes constant.

With these equations the questions on asymmetry posed above can be answered (about the locus of the flankers in the retinotopic map and whether asymmetry is compatible with that): The locus of the crowding target is given by Equation 12 and the distance of the flankers from that location is given by Equation 13. For small eccentricities—which we are interested here—the cortical spacing depends on the precise $E_2$ values and is difficult to predict. Beyond a certain eccentricity ($>3^\circ$), however, cortical critical spacing quickly becomes constant as Pelli (2008) has derived. If we apply these rules to the asymmetry in the confusion rate found here (Figures 7 and 8), it follows that with a symmetrical arrangement of flankers as used in a typical crowding task (for which $d$ is equal for the left and right flanker and thus not proportional to $E$) the corresponding distances on the primary visual cortical surface are asymmetrical and thus would suggest asymmetrical flanker interference. However, whether at small eccentricities that asymmetry is in accord or opposite to the empirically found, is again difficult to decide. For sufficiently large eccentricity the cortical spacing for the outer flanker is smaller and the prediction is that of more interference from the outer flanker. This is what is found for feature interactions (e.g., Bouma, 1970; Estes & Wolford, 1971). It is opposite to that found for letter-source confusions (Chastain, 1982, and present results) which seem to play an opposite role. It will thus be interesting to further study how the asymmetries vary with eccentricity, particularly at positions close to the fovea, and whether the different contributions discussed above can be further disentangled.
Letter source confusions are a major cause of crowding. A transient positional cue appearing shortly before the target attracts transient spatial attention and reduces crowding. The way, however, by which transient attention exerts its effect, is not by influencing letter-source confusion but rather by improving contrast processing. Whether this happens by decreasing feature-source confusion (which is widely held to underlie crowding), by decreased surround suppression, or by some other way we cannot say. The reverse asymmetry of flanker confusions compared with feature mislocalization could mean, though, that the two are complementary mechanisms, both contributing independently. In effect, transient attention improved contrast gain by a factor of up to 1.7 relative to its surroundings. It is as if the metaphorical spotlight (dating back to von Helmholtz; cf. van der Heijden, 1992, p. 33) shines on a retinotopic area, perhaps V1 or the lateral geniculate nucleus (Hillyard, Vogel, & Luck, 1992, p. 33) shines on a retinotopic area, perhaps V1 or the lateral geniculate nucleus (Hillyard, Vogel, & Luck, 1992; Fang & He, 2008; Saalmann & Kastner, 2009), to improve its processing and/or suppress the processing of the surround. Note that this account is neuro-analytic, as requested by Tyler and Likova (2007), since gain change can be implemented neurocomputationally by changed synaptic coupling from multiplicative back projections onto an earlier retinotopic area (like V1, CGL, pulvinar, thalamic nucleus reticularis, etc.). Surprisingly, gain improves independently of the ring cue’s size—only its center position matters.

Gain control by the ring cue

Figure 10 summarizes and visualizes the gain-control effect of transient attention on contrast processing (as observed in the crowding situation, i.e., constrained by the crowding-induced loss of target contrast sensitivity) in a semiquantitative way. The gray values of the transparent mask in the graph follow, along the four half-meridians, the effect’s shape from Figure 4, drawn to scale; i.e., they show the reduction of the flankers’ effect on target contrast threshold by a ring cue (of unimportant size), centered on the target, as a function of flanker distance. Suppression is maximal not far from the target, leaving a free space between target and flanker of 12.5% of the target’s eccentricity E (Equation 6). Towards the center, suppression decreases steeply. Outwards from the maximum, the effect decreases gently and linearly and vanishes at a distance when there is 70% of the target’s eccentricity free space to the target (cf. Equation 5, which describes the vanishing point in Figure 4; note that the outer limit is constrained by, but not equal to, critical crowding distance, as explained in Results). Expressed as center-to-center distance, the proportions are a little higher, the maximum being roughly at 20% E, and the outer bound at 80% E. What the 2D geometrical shape of the suppression field is (whether it is circular or elliptic) we cannot say since we measured on the horizontal meridian only. From Toet and Levi (1992) we expect the form to be elliptic with the main axis oriented radially.

Ring cue’s effect on confusions

Unlike contrast processing, letter source confusion turns out to be independent of the transient cue but takes place in about the same area around the target. With a linear percent-confusion scale (compared to a log contrast scale), the shape of the effect size is also rather similar (Figure 6). Confusions are highest—i.e., position code weakest—at a flanker separation that leaves 18.8% of E free space to the target (Equation 8). Perhaps this is the space required so that flanker features remain bound to the flanker rather than floating to the target. At the maximum, nearly 40% of misidentifications stem from “character jumbling,” which leaves 60% of recognition errors to be due to other causes. From there inward, towards the target, confusions decrease steeply. Away from the target, confusions decrease gently and linearly until they vanish when there is 62.5% of E free space to the target (Equation 7).

Conclusion

In summary, out of Korte’s seven Gestalt principles of form perception in indirect vision, our data document the role of one, false localization of letters. We term this phenomenon letter source confusion. It is one of the main contributors to the crowding effect, along with altered contrast processing and false localization of detail. Transient spatial attention, as assessed by the effects of a transient ring cue, substantially reduced the crowding effect by increasing target contrast sensitivity relative to the standard crowded display, or by suppressing the effective contrast of the flankers. The induced contrast gain depended on flanker distance; at its maximum it amounted to a factor of 1.7. The size of that ring cue seemed irrelevant for its effect, and the area in which transient attention operated was governed by visual field location only. That area can be described as a ring around the target that fills the area in which crowding takes place and scaled very precisely with the target’s...
eccentricity. We can thus generalize Bouma’s critical-distance rule to describe the free space between flankers and target that leads to maximum or to vanishing cue effect in the crowding situation. The proportions were 12.5% and 70% of target eccentricity, respectively. For flanker distances expressed as center-to-center (as now more common), the corresponding values in the generalized Bouma rule were roughly 20%/80% of eccentricity. Bouma’s rule shows a formal, and perhaps causal, analogy to $M$ scaling, from which the critical crowding distances on a cortical map can be derived as a logarithmic function.

False localization of letters was independent of transient attention, i.e., not influenced by a transient cue. This is surprising and shows that localization on the letter level is not helped by a transient positional cue. These false localizations occurred in that same ring area around the target as the attentional effect and depended similarly on eccentricity in extent. At its maximum, letter source confusion$^3$ accounted for nearly 40% of recognition errors. The maximum was close to where enhancement by the cue was maximal, namely at 19% of target eccentricity free space left and right of the target. Further away from the target, letter source confusion vanished at 62.5% free space to the target. False localizations further showed anisotropy, with the flanker inward in the visual field being more easily confused with the target. This asymmetry increased linearly with increasing eccentricity while the outward-flanker confusion rate stayed constant. The asymmetry is opposite to that reported for featural interference. In summary, our results are a psychophysical counterpart to a separate neural coding of what and where in pattern recognition. The results can be summarized in a “doughnut” Bouma rule.

**Acknowledgments**

Maka Malania was supported by the German Academic Exchange Service (DAAD), Refr. # A/07/04655. Travel to the ECVP was supported by the Deutsche Forschungsgemeinschaft, grant STR 354/7-1 to HS; open access costs were borne by the University of Göttingen. We thank Denis Pelli for fruitful discussions, Manfred MacKeben, Ronald van den Berg, and four reviewers for helpful comments on the manuscript, Josh Solomon for bringing up asymmetry, a reviewer pointing out possible distinctions between cortical mapping and crowding asymmetries, Neil Parry for English corrections, and Barbara Herzberger for translating the Korte excerpt.

Commercial relationships: none.

**Footnotes**

1 “When one is testing amblyopic children with isolated letters or E’s, the visual acuity recorded is often much better than with the ordinary test chart. If the visual field is crowded with letters, the area of the visual field in which the letters can be recognized narrows. This is very easy to demonstrate, as I showed at the Congress of Scandinavian Ophthalmologists in 1936.”

2 Terminology: In our own work we have mostly used digits, not letters, as stimuli and spoke of character contrast thresholds. However, the term letter contrast threshold is well established now, and since there seem to be no systematic differences quantitatively we use the more common term.

3 Further on terminology: In a previous abstract (Strasburger & Malania, 2008) we had used the term “source confusion” in the sense that an entire character at a different position is taken for the target. That term “source confusion” has, however, already been used in a different meaning in the context of featural accounts of letter recognition (e.g., Huber, Shiffrin, Lyle, & Quach, 2002; Nandy & Tjan, 2007, p. 2). That meaning implies a mechanism where letter confusability (Wolford, 1975; Krumhansl & Thomas, 1977) arises from letters sharing certain features; the respective source of each of those features is from a different location and leads to the confusion. Strasburger et al. (2011) therefore distinguished the two kinds of source confusion for clarity and we use that terminology here.

**References**


Harvey, L. O., Jr. (1986). Efficient estimation of...


