



# Psychophysics of reading XX. Linking letter recognition to reading speed in central and peripheral vision

Gordon E. Legge<sup>a,\*</sup>, J. Stephen Mansfield<sup>a,1</sup>, Susana T.L. Chung<sup>b</sup>

<sup>a</sup> Minnesota Laboratory for Low-Vision Research, University of Minnesota, 75 East River Rd, Minneapolis, MN 55455, USA

<sup>b</sup> School of Optometry, Indiana University, 800 East Atwater Ave, Bloomington, IN 47405, USA

Received 27 March 2000; received in revised form 30 August 2000; accepted 16 November 2000

## Abstract

Our goal is to link spatial and temporal properties of letter recognition to reading speed for text viewed centrally or in peripheral vision. We propose that the size of the visual span — the number of letters recognizable in a glance — imposes a fundamental limit on reading speed, and that shrinkage of the visual span in peripheral vision accounts for slower peripheral reading. In Experiment 1, we estimated the size of the visual span in the lower visual field by measuring RSVP (rapid serial visual presentation) reading times as a function of word length. The size of the visual span decreased from at least 10 letters in central vision to 1.7 letters at 15° eccentricity, in good agreement with the corresponding reduction of reading speed measured by Chung and coworkers (Chung, S. T. L., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research*, 38, 2949–2962). In Exp. 2, we measured letter recognition for trigrams (random strings of three letters) as a function of their position on horizontal lines passing through fixation (central vision) or displaced downward into the lower visual field (5, 10 and 20°). We also varied trigram presentation time. We used these data to construct visual-span profiles of letter accuracy versus letter position. These profiles were used as input to a parameter-free model whose output was RSVP reading speed. A version of this model containing a simple lexical-matching rule accounted for RSVP reading speed in central vision. Failure of this version of the model in peripheral vision indicated that people rely more on lexical inference to support peripheral reading. We conclude that spatiotemporal characteristics of the visual span limit RSVP reading speed in central vision, and that shrinkage of the visual span results in slower reading in peripheral vision. © 2001 Elsevier Science Ltd. All rights reserved.

*Keywords:* Reading; Letter recognition; Peripheral vision; Low vision; Visual span

## 1. Introduction

Reading speed has been widely studied, but its relationship to letter recognition remains unclear. Our goal is to establish this link for both central and peripheral vision.

### 1.1. Visual span in normal vision

Although people have the strong impression of seeing

a whole page of text simultaneously, it has long been known that only a few letters are recognized on each fixation. We propose that reading speed is limited by the number of letters that can be recognized in parallel. Following O'Regan (1990, 1991), we call this the 'visual span.' He defined the visual span as the region around the point of fixation within which characters of a given size can be resolved<sup>2</sup>. The boundary of the visual span

\* Corresponding author. Tel.: +1-612-6250846; fax: +1-612-6262079.

E-mail address: legge@eye.psych.umn.edu (G.E. Legge).

<sup>1</sup> Present address: Department of Psychology, Plattsburgh State University, 101 Broad Street, Plattsburgh, NY 12901, USA.

<sup>2</sup> The notion of 'visual span' differs from the concept of 'perceptual span' (McConkie & Rayner, 1975). The size of the 'perceptual span' depends on factors in addition to letter recognition. Operationally, it refers to the region of visual field that influences eye movements and fixation times in reading. Rayner and McConkie (1976) estimated that the perceptual span extends 15 characters to the right of fixation and four characters to the left.

indicates the horizontal retinal eccentricity at which letters, formatted as in text, are no longer recognizable in reading. This boundary is jointly determined by decreasing letter acuity in peripheral vision, and lateral masking (crowding) between adjacent letters (Bouma, 1970). A consequence of the linear scaling laws that apply to both peripheral letter acuity and crowding (cf. Wilson, Levi, Maffei, Rovamo, & DeValois, 1990) is that the size of the visual span is roughly constant when measured in letter spaces over a moderate range of angular character size. This scale invariance means that one cannot squeeze more letters into the visual span by using smaller print.

O'Regan (1990, 1991) presented a theoretical model of the visual span in central vision based on the size of the critical features in the letters, the fall-off in the eye's spatial resolution away from the fixation point, and the geometry of the display surface. His model predicts a visual span of about 15 characters for letters subtending  $0.4^\circ$ . O'Regan, Levy-Schoen, and Jacobs (1983) measured the recognition of letters (flanked by numerals) as a function of their retinal eccentricity. They defined visual span in terms of the eccentricity within which letters could be recognized above some criterion level. For criteria of 50 and 90% correct, the visual spans were 22 letters and 10 letters, respectively. This finding makes clear that the numerical size of the visual span is criterion-dependent.

Rayner and Bertera (1979) used an eye-tracking method to mask letters (each subtending about  $0.33^\circ$ ) surrounding the point of fixation during reading. When the mask covered the central seven letters, reading speed was very low, about 12 words/min. When the mask covered 11 letters, reading was essentially impossible. Fine and Rubin (1999) showed that these masking effects scale with letter size. These results imply that human readers have a visual span of 7–11 letters.

Legge, Ahn, Klitz, and Luebker (1997a) measured reading times as a function of word length using a rapid-serial visual-presentation (RSVP) method. From their data, they estimated the size of the visual span for high-contrast  $1^\circ$  letters to be 10.6 letters.

These empirical estimates point to a visual span in normal central vision of about 10 letters over a moderate range of print sizes<sup>3</sup> ( $\sim 0.3$ – $1.0^\circ$ ). Nobody, to our knowledge, has obtained empirical estimates of the size of the visual span in peripheral vision. Nor has anyone explained how the size of the visual span influences reading speed.

<sup>3</sup> There is both theoretical evidence (O'Regan, 1991) and empirical evidence (Legge et al., 1997a) that the visual span is smaller for highly magnified text ( $6^\circ$  letters).

## 1.2. *Visual span in low vision*

Shrinkage of the visual span may play an important role in explaining reduced reading speed in low vision.

Some people with low vision have reduced retinal–image contrast, resulting from cloudy ocular media, or reduced contrast sensitivity with a neural basis. These individuals might be expected to have reduced visual spans, based on findings from normal subjects. Legge et al. (1997a) estimated the visual spans of normal subjects to decrease from 10.6 to 1.7 characters as text contrast decreased from 100 to 1.5%. They also obtained estimated visual spans for a group of seven low-vision subjects. The low-vision spans ranged from normal values to less than one character.

These findings fit the following interpretation. When retinal–image contrast (or, equivalently, contrast sensitivity) drops, the visual span for reading gets narrower, and the reader recognizes fewer letters per fixation. As a consequence, the reader saccades through text in smaller steps with a corresponding reduction in reading speed.

The most common cause of low vision in developed countries is macular degeneration, which often results in a scotoma in central vision. People with central scotomas usually have severe reading difficulty even when adequate magnification is provided to compensate for their acuity deficits (Faye, 1984; Legge, Rubin, Pelli, & Schleske, 1985; Legge, Ross, Isenberg, & LaMay, 1992; Whittaker & Lovie-Kitchin, 1993). Explaining the reading deficits of people with central scotomas remains an important unresolved problem.

People with central scotomas must use peripheral vision for reading. Shrinkage of the size of the visual span in peripheral vision could impose a bottleneck on reading performance for these individuals.

Measurements of reduced saccade lengths in readers with central scotomas are consistent with a smaller visual span (Rumney & Leat, 1994; Bullimore & Bailey, 1995).

## 1.3. *Linking the visual span to reading speed*

Chung, Mansfield, and Legge (1998) measured reading speed as a function of print size in peripheral vision for six normally sighted subjects. Following Rubin and Turano (1994), they used the RSVP method, rather than conventional page reading, so that text was confined to a localized region of the retina, and to reduce the need for eye movements. The stimuli were short sentences, presented at six eccentricities from 0 to  $20^\circ$  in the lower visual field.

At all eccentricities, reading speed rose with print size until a critical print size was reached, and then leveled out at a maximum value. These maximum reading speeds varied with retinal eccentricity, dropping by

about a factor of six from central vision (862.1 wpm) to 20° eccentricity (142.6 wpm).

What is the relationship between reading speed and the size of the visual span? Can this relationship explain slower reading in peripheral vision?

Legge, Klitz, and Tjan (1997b) studied the performance of an ideal-observer model of reading, implemented as a computer simulation named *Mr. Chips*. The size of the visual span is a key parameter of the model. Simulation results showed that *Mr. Chips*'s mean saccade length  $M$  varied with the number of characters  $N$  in the visual span according to the simple equation  $M = N + 1$ . Assuming that reading speed is proportional to mean saccade size, this equation predicts a five-fold reduction in reading speed when the visual span decreases from nine characters to one character.

Although this ideal-observer model makes a direct prediction about the relationship between visual-span size and saccade size, we must address three major issues in bridging the gap to human reading:

1. Temporal dependence of the visual span: Because accuracy of human letter recognition depends on exposure time, we would expect the size of the visual span to be time-dependent. The *Mr. Chips* model does not explicitly address the role of temporal factors.
2. Letter-Recognition Errors: The *Mr. Chips* model and many others (e.g., Clark & O'Regan, 1999) assume flawless letter recognition within the visual span. But empirical measurements of the human visual span (O'Regan et al., 1983; Nazir, O'Regan, & Jacobs, 1991) show profiles of decreasing accuracy. In Exp. 2, we will display data showing how letter-recognition accuracy depends on exposure time and spatial position. We will address the issue of how letter-recognition errors can be handled in Section 4 of this paper.
3. Lexical Inference: *Mr. Chips* uses lexical knowledge to identify some words on the basis of partial information. For example, *Mr. Chips* would infer that 'differ???' must be 'different' if there is no other nine-letter word beginning with 'differ' in the model's lexicon. In the same spirit, Clark and O'Regan (1999) have described a model that recognizes only the two letters of words nearest fixation and the end letters, but takes advantage of orthographic constraints and lexical inference in word recognition. Both the *Mr. Chips* analysis and the Clark and O'Regan analysis suggest a strong role for lexical inference in rapid word recognition and reading. On the other hand, results of theoretical modeling prompted Legge et al. (1997b) to speculate that human readers may sacrifice lexical inference for computational sim-

licity in rapid reading. In section 4 on modeling, we will ask if human letter-recognition data by itself is sufficient to account for the empirical RSVP reading speeds or whether lexical inference also plays a role.

#### 1.4. Plan of this paper

Our interest in the visual span is two-fold. From a theoretical point of view, the visual span can provide a way of characterizing the front-end (bottom-up) visual information available to the reading process. From a clinical point of view, the visual span may be a theoretical construct that helps explain why reading slows down when retinal image contrast is low, and when there are central scotomas.

In Exp. 1, we used the Legge et al. (1997a) method for estimating the visual span in peripheral vision. The results showed that the reduction in visual span does parallel the reduction in reading rate, in qualitative agreement with the Chung et al. (1998) reading-speed data, but individual differences led us to worry about contamination of the results from top-down strategies.

In Exp. 2, we used a letter-recognition method, avoiding the top-down influences in Exp. 1. We used these data to construct visual-span profiles (letter-recognition accuracy vs. letter position) in central and peripheral vision.

Finally, we will present a model that takes the letter-recognition data from Exp. 2 as input and produces RSVP reading speeds as output. We will show that the letter-recognition data, together with the simplest possible lexical-matching operation, are sufficient to account for RSVP reading speed in central vision. In peripheral vision, comparison of data and model indicate that people invoke some form of lexical inference to increase reading speed.

## 2. Experiment 1. Estimating the visual span using RSVP reading

We tested the *shrinking visual span hypothesis*; the number of letters that are recognized on each glance shrinks in peripheral vision. If so, more time should be needed to recognize words whose lengths exceed the size of the visual span because two or more glances would be necessary. An indicator of a shrinking visual span in peripheral vision would be an increased dependence of word-recognition time on word length. We used the method developed by Legge et al. (1997a). This method is similar to one used by Farah and Wallace (1991) in studying acquired dyslexia.

## 2.1. Method

### 2.1.1. Stimuli

Words were presented one after the other at the same place on the screen (RSVP), at a specified eccentricity. The words were rendered in Courier Bold and were displayed as dark letters on a white (110 cd/m<sup>2</sup>) background with a contrast of ca. 90%.

We tested five retinal eccentricities — 0° (central vision), and 2.5, 5, 10 and 15° in the lower visual field. At each eccentricity, the character size was chosen to be twice as large as the corresponding mean critical print size (CPS) measured by Chung et al. (1998)<sup>4</sup>. Table 1 lists the log MAR print sizes, corresponding min-arc values, and viewing distances for the five retinal eccentricities. By scaling letter size in peripheral vision, we tested performance under conditions in which reading speed is independent of print size. Although we tested only one print size at a given eccentricity in this study, we expect that our visual-span results would also be independent of print size (for values exceeding the CPS).

### 2.1.2. Procedure

At each eccentricity, we measured RSVP reading speed as a function of word length from four to 10 letters.

For eccentric stimuli, subjects were instructed to maintain their fixation along a horizontal red line that

Table 1  
Print sizes in Experiments 1 and 2

Eccentricity	log MAR	min-arc (x-height)	Viewing distance (cm)
<i>Exp. 1</i>			
0°	0.58	19	120
2.5°	1.04	55	40
5°	1.22	83	40
10°	1.51	162	40
15°	1.58	190	40
<i>Exp. 2</i>			
0°	0.78	30	40
5°	1.38	120	40
10°	1.62	210	40
20°	1.78	300	30

<sup>4</sup> ‘Critical print size’ (CPS) is defined to be the smallest print size that yields maximum reading speed. CPS is defined for a line of text, not for a strictly local point on the retina. It is a functional measure, indicating when reading speed no longer depends on print size, not a local property of the visual field. When we refer to the CPS at a given eccentricity, we are referring to a horizontal line of text that intersects the vertical midline at this retinal eccentricity. From an evaluation of the mean CPS values in peripheral vision measured by Chung et al., and the variability across subjects, we selected print sizes in the present study which should exceed CPS values for all subjects — 2 × CPS in Exp. 1, and 2.5 × CPS in Exp. 2.

paralleled the words. They were permitted to make horizontal eye movements, but not vertical eye movements. We used a video eye tracker (ISCAN RK 416, Boston, MA) to ensure that fixation did not drift below the red line. Prior to a block of RSVP trials, we measured the subject’s fixation for 1 min to obtain an average vertical eye position. If the vertical eye position during the trial deviated from the ‘average’ by two standard deviations, the trial was rejected. 2.1% of our trials were rejected by this criterion.

Subjects initiated each trial by pressing a key. A trial consisted of a series of six words of fixed length. The six words were preceded and followed by strings of *x*s equal to the word length. In each trial, words were drawn at random, without replacement, from pools of 750 words for each length. The 750 most common 10-letter words were used, and the frequencies of words in pools of other lengths were matched to these (Kilgarriff, 1997).

For words of a given length, we used a staircase method to cluster trials near the 80%-correct exposure time. Each staircase had 20 trials. Exposure time was increased by 0.1 log units if more than one word was read incorrectly, kept the same if exactly one error was made, and reduced by 0.1 log units if no errors were made.

Psychometric functions, percent correct versus log RSVP exposure time, were created by fitting these data with cumulative Gaussians of a fixed slope of 3.0. This slope is the average of the slopes we obtained from pilot data, and is also close to the average slope for psychometric functions obtained by Chung et al. (1998). The threshold exposure time, for words of a given length was based on the 80% correct point on the psychometric function. Error bars for these threshold exposure times were computed using a parametric bootstrapping method (Foster & Bischof, 1991), and these error bars were subsequently used in fitting straight lines to the threshold-time versus word-length data. Each staircase was run two times for a total of 40 trials per psychometric function (or 240 words per reading-speed threshold estimate). The staircases (one for each length) were randomly interleaved within each block.

### 2.1.3. Subjects

There were eight normally sighted subjects, with corrected acuities ranging from 20/10 to 20/20. To minimize optical aberrations in eccentric viewing, we avoided subjects with spectacle corrections. All the subjects were either emmetropic or wore contact lenses. None of our contact-lens wearers had astigmatism that required correction using toric lenses, which might also introduce aberrations in peripheral vision.

All subjects had practice sessions prior to data collection. Only subjects F and G had participated in a previous study using peripheral viewing. Informed con-

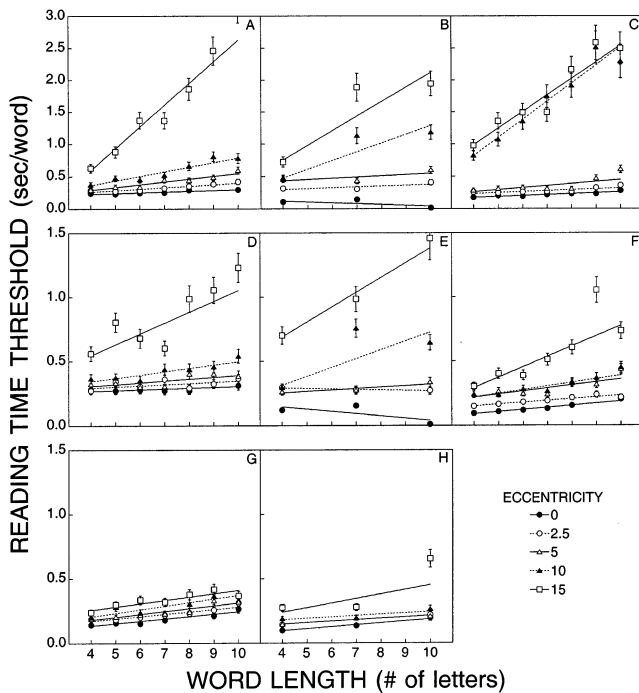


Fig. 1. Reading-time thresholds for words (80% accuracy criterion) are plotted as a function of word length for the eight subjects in Exp. 1. Data are shown for five eccentricities from 0 to 15° in the lower visual field. Best-fitting straight lines are shown for each data set.

sent was obtained from all subjects prior to participation.

## 2.2. Results

Fig. 1 shows reading-time thresholds as a function of word length, one panel for each of the eight subjects. There are separate data sets in each panel for the five retinal eccentricities, each summarized by a linear regression line:

$$T = a + bL$$

where  $T$  is reading time,  $L$  is word length, and  $a$  and  $b$  are eccentricity-dependent intercept and slope parameters.

Shrinkage of the visual span in peripheral vision can be tested with three statistical hypotheses: (1) the slopes,  $b$ , should increase with increasing retinal eccentricity, (2) the intercepts,  $a$ , should be approximately equal across eccentricities<sup>5</sup>, and (3) there should be a statistically significant interaction effect on reading time between word length and retinal eccentricity.

<sup>5</sup> As words become very short, the effect of eccentricity should diminish because the words will fit entirely within a small visual span. This leads to the prediction of no intercept variation. Growth of the intercept with increasing eccentricity would suggest an alternative model in which the visual span remains large in peripheral vision but reduced reading speed results from much slower processing within a constant visual span. This 'prolonged viewing' alternative is discussed in more detail by Legge et al. (1997a).

All three predictions are confirmed by statistical tests. Consistent with the first two predictions, analyses of variance (one-way ANOVA of parameter estimate,  $a$  or  $b$ , versus eccentricity for the eight subjects) show that there was no significant effect of eccentricity on the intercept  $a$  [ $F(4,28) = 1.305$ ,  $P = 0.292$ ], but there was a significant effect of eccentricity on the slope  $b$  [ $F(4,28) = 5.94$ ,  $P = 0.001$ ]. To test the third prediction, we did a repeated-measures analysis of variance on reading time as a function of word length and retinal eccentricity. As predicted, there was a significant interaction between word length and eccentricity [ $F(24,168) = 4.92$ ,  $P < 0.001$ ]. If the visual span in central vision (0° eccentricity) is wide enough for 10 or more characters to be identified in parallel, regression lines through reading time versus word length should be flat with slopes close to zero. The curves of all eight subjects for 0 and 2.5° were nearly flat, with slopes close to zero.

If the visual span is much narrower for eccentric viewing, the reader must proceed in smaller steps, recognizing only a few letters at a time. If this is the case, the stronger dependence of reading time on word length will be manifest in an increase in the slope. Except for the aberrant subject G, regression lines for all subjects get steeper in the periphery, indicative of a shrinking visual span.

We might expect the shrinking visual span to yield curves in Fig. 1 that are stair steps, rather than straight lines, with the distance between steps being related to the size of the visual span. But variation in fixation location within words and graded profiles rather than sharp edges of the visual span (see Exp. 2) would 'blur out' the stair steps, and account for the straight lines we observed.

We can estimate the size of the visual span as follows. The slopes  $b$  of the regression lines in Fig. 1 have units of time per letter (i.e. time per word divided by the number of letters per word). The reciprocal slope,  $1/b$ , is the number of letters identified per unit time. If we assume that the reading task consists of a series of glances (fixations) of approximately equal duration, the reciprocal slope can be used to estimate the number of letters identified per glance.

The average fixation time in reading is about 250 ms (Rayner & McConkie, 1976). Accordingly, for Exp. 1, we operationally define *visual span* to be equal to  $250/b$  where  $b$  is the slope from regression lines of reading time (ms) versus word length (as in Fig. 1).

Fig. 2 shows average slopes (left vertical scale) and corresponding visual spans (right vertical scale) versus retinal eccentricity for our eight subjects. The shading indicates a ceiling of 10 letters on our estimates of the size of the visual span, since our longest target words were 10 letters. The large error bars reflect the individual differences. For 10 and 15° eccentricities, there is clear

evidence for a shrinking visual span. The average visual span estimate at  $10^\circ$  is 3.08 letters and at  $15^\circ$  it is 1.71 letters.

### 2.3. Discussion

These findings are consistent with the following interpretation. When text is presented in peripheral vision, the visual span gets narrower; the reader recognizes fewer letters per fixation. As a consequence, the reader saccades through the text in smaller steps with a corresponding reduction in reading speed. In effect, the capacity of normal vision to process several letters in parallel deteriorates in peripheral vision. According to this interpretation, a fundamental limit to reading in peripheral vision is the reduction in the size of the visual span. This may explain the reading deficits commonly observed in people with central scotomas.

Chung et al. (1998) found a 4.4-fold reduction in maximum reading speed at  $15^\circ$  eccentricity compared with central vision. The estimated reduction of the visual span in the present study from at least 10 in central vision to 1.71 letters at  $15^\circ$  (a factor of 5.8) is about 30% larger.

Correction of our estimates of the visual span would be necessary if mean fixation times differ substantially from 250 ms. Exp. 3 in Legge et al. (1997a; Exp.3) recorded fixation times and made these corrections in estimating visual spans for text with different contrasts. Original and corrected values were 10.6 and 9.3 characters for 100% contrast, and 1.74 and 2.83 characters for 1.5% contrast. There is evidence for a modest increase in

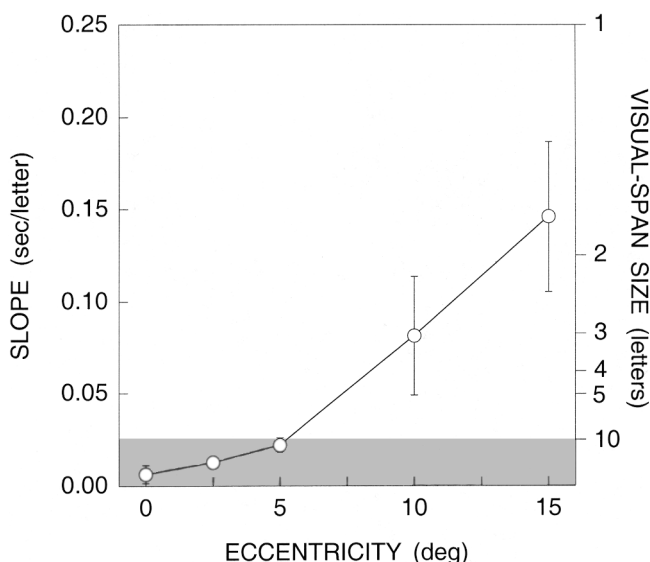


Fig. 2. Mean slopes across the eight subjects in Fig. 1 are plotted as a function of eccentricity. The right vertical scale shows the corresponding estimates of visual span (see text for derivation). The shaded region indicates that we could not estimate visual spans larger than 10 in Exp. 1 because the longest words were 10 letters in length.

fixation times for reading in peripheral vision. Fine and Rubin (1999) measured fixation times of about 320–350 ms for normal subjects with simulated central scotomas. Trauzettel-Klosinski, Teschner, Tornow, and Zrenner (1994) compared reading eye movements in normal subjects and patients with central scotomas resulting from juvenile forms of macular degeneration. Mean fixation time for the patients was 36% longer than for the normal subjects (our estimate, inferred from mean reading speeds and saccade sizes in their Fig. 6). Bulimore and Bailey (1995) compared reading eye movements in normal subjects and patients with age-related macular degeneration, but found no important differences in fixation rates. From these studies, we estimate that prolonged fixation times in peripheral vision would necessitate, at most, correction of our peripheral visual-span estimates by a factor of 1.5. Exp. 2 will provide a more direct assessment of temporal influences on the visual span.

Two aspects of Experiment 1 leave the concept of the visual span uncertain, one empirical and one theoretical.

Empirically, the data of Fig. 1 show substantial individual differences. For instance, subject G is aberrant in that even the peripheral slopes are quite shallow. Although we cannot be sure of the source of these differences, informal debriefing of the subjects suggests that they used different eye-movement or lexical-inference strategies for dealing with targets in peripheral vision.

These individual differences raise the possibility that our estimates of the visual span were influenced by higher-level strategies. These might include oculomotor strategies for planning and executing saccades to stimuli in peripheral vision, or lexical-inference strategies for ‘guessing’ words from sparse visual data. These strategic influences are undesirable because we conceive of the visual span as a bottom-up sensory limitation on the capacity to recognize letters. Exp. 2 uses a method for measuring the visual span that is free of oculomotor or lexical influences.

From a theoretical point of view, estimates of the size of the visual span do not, by themselves, provide a prediction for reading speed. The similarity in drop-off between reading speeds (Chung et al., 1998) and visual-span estimates in peripheral vision (Fig. 2) is promising, but does not explain the relationship between these quantities. Following the description of Exp. 2, we will outline a model that links letter recognition within the visual span to reading speed.

### 3. Experiment 2. Measuring spatial profiles of the visual span

We developed a second method for measuring the visual span that (1) explicitly takes into account expo-

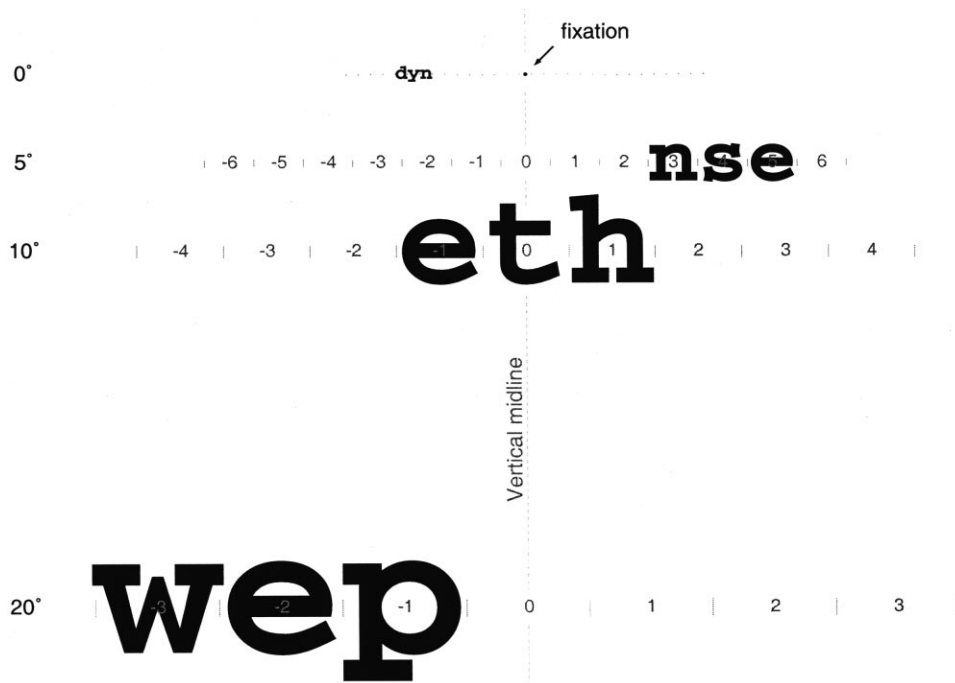


Fig. 3. Examples of trigram stimuli in Exp. 2. Trigrams were presented on four horizontal lines intersecting the vertical midline at 0, 5, 10 and 20° in the lower visual field. Trigrams were scaled in size to exceed the critical print size for the corresponding eccentricity. Position left or right of the midline was specified in units of letter size. All letters were lowercase Courier Bold. The viewing distance should be  $2.75 \times$  the figure height for the trigrams to appear at the retinal locations and size used in Exp. 2.

sure time; (2) does not depend on eye-movement strategies; and (3) does not depend on lexical inference. Our goal was to study the visual span in a more direct way by measuring letter recognition per se, and to minimize opportunities for individual differences due to top-down strategies.

### 3.1. Methods

#### 3.1.1. Subjects

There were three subjects. Their binocular acuities were 20/15 (TAH and STC) and 20/20 (JSM). The subjects were tested binocularly. Two were authors and had participated in previous studies of peripheral vision. The third (TAH) was a member of the lab staff not familiar with the details of the project, and had no prior experience with peripheral vision experiments. All subjects had practice sessions before data collection. This experiment involves letter recognition in peripheral vision. Although some types of visual performance improve with practice in peripheral vision, such as vernier acuity (Beard, Levi, & Reich, 1995), recent evidence indicates that practice has little or no effect on acuity measured with Landolt C's (Westheimer, 1999).

#### 3.1.2. Stimuli

The targets were trigrams, random strings of three letters. As shown in Fig. 3, we measured performance for trigrams on four horizontal lines in the visual field

— 0° passing through the point of fixation, and 5, 10, and 20° in the lower visual field. Along each of these lines, positions are indicated by the number of letter slots left (negative values) or right (positive values) of the midline.

At all eccentricities, we scaled the letter size to about 2.5 times the critical print size (Chung et al., 1998). Values are given in Table 1.

As in Exp. 1, the letters were rendered in Courier Bold and were displayed as dark letters on a white (110 cd/m<sup>2</sup>) background with a contrast of ca. 90%. The character spacing within the trigrams was the normal spacing in the Courier-Bold font<sup>6</sup>.

#### 3.1.3. Procedure

Trigrams were composed of a random selection of three letters (repeats were possible). They were presented for exposure times of 25, 40, 50, 80, 125, 200, 300 and 500 ms. The central letter of the trigram could occupy letter position 0 (on the midline), or positions 1 to 10 right or left from the midline position. (For the more eccentric conditions, the larger letter sizes used required that the range of letter positions was reduced, as indicated in the data plots.)

<sup>6</sup> If  $x$  is letter size (x-height) in min-arc, as in Table 1, and  $n$  is the number of the letter positions left or right of midline (0 for the midline letter), then a letter in the  $n$ th position is centered at  $1.1 \times x$  min-arc horizontally from the midline.

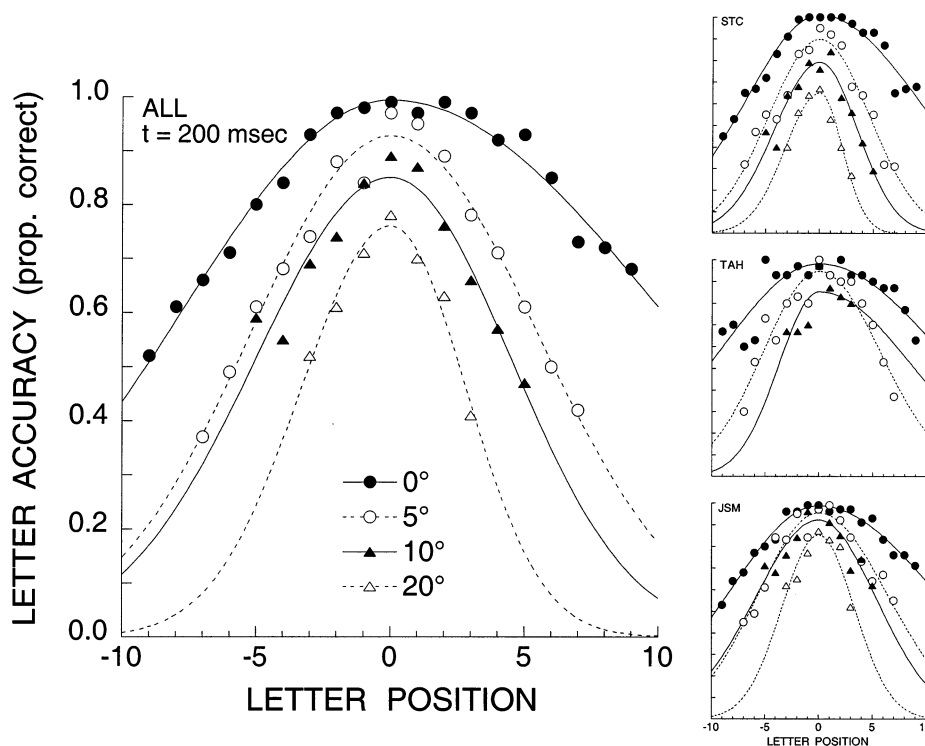


Fig. 4. Effect of eccentricity. Trigram recognition data from Exp. 2 are shown for three individual subjects (STC, TAH, and JSM) and for the group average (ALL). Letter recognition accuracy (proportion correct) is plotted as a function of letter position left or right of the midline. Data at the four eccentricities have been fit with split Gaussians (see text). The resulting curves are termed *visual-span profiles*. The exposure time was 200 ms.

Subjects were instructed to fixate on a small red fixation dot throughout a trial for all peripheral testing. For foveal testing, the red fixation dot appeared prior to a trial, but disappeared during the trigram presentation. Eye movements were not monitored, but the subjects had all demonstrated good fixation in previous eye-tracking measurements.

Each trial began with the string # # # presented on the monitor to cue the subject to the position of the trigram. Following presentation of the trigram, the subject was required to identify the three letters verbally<sup>7</sup>.

In each block of trials, we tested a selection of trigram positions. Trigram positions were chosen such that adjacent letter positions were tested in different blocks. All the exposure durations were tested for each letter position within the same block. Each condition

(letter position by exposure duration) was tested 10 times in the same block. Trials were repeated (at a randomly-determined later time in the block) when the subjects indicated that they had unintentionally blinked or failed to maintain fixation. Because the number of letter positions tested varied between blocks, the number of trials per block varied from 90 to 630 trigrams (mean = 300).

The subjects reported all three letters of the trigram from left to right. For a letter to be correct, it had to be reported in the correct position.

The data were used to construct plots of letter-recognition accuracy (proportion correct) versus horizontal letter position (see examples in Figs. 4–6). These plots were fit with split Gaussians characterized by three parameters: the amplitude, and the standard deviations of the left and right sides. The curve fits accounted for 85–95% of the variance for different plots. These curves are spatial profiles that characterize the visual span.

In a control experiment, two subjects (JSM and STC) were tested in a masking condition. A mask consisting of the symbols # # # immediately followed the trigram and remained on the screen until the subject made a response. Masked visual-span profiles were measured at 10° eccentricity for exposure times of 40 and 200 ms. For the other eccentricities, effects of masking were assessed only along the midline.

<sup>7</sup> By chance, some of the trigrams were three-letter English words. To assess the impact of these word trigrams on performance, we compared letter recognition accuracy on 290 trials on which this occurred (1.8% of the total number of trials), with 290 non-word trials matched for stimulus conditions but otherwise randomly selected. Recognition accuracy was 66.2% for the word trigrams and 64.1% for the 'non-word' trigrams. Further analysis of 290 unpronounceable trigrams (composed of three consonants) yielded 63.2% correct. Given these small performance differences, we believe that the linguistic properties of the stimuli had little impact on our results.



### 3.2. Results

In Fig. 4, the large panel shows group data, and the three smaller panels individual data. Letter accuracy (proportion correct) is plotted as a function of letter position left or right of midline, with separate curves for each eccentricity. The presentation time was 200 ms. Data points for the individual subjects each represent 30 trials at a given letter position. These points combine data across cases in which the letter position contained a middle, left or right component of the trigram.

All the curves peak near 0 (midline) and are slightly broader on the right than left. (The asymmetry will be discussed in more detail below in connection with Fig. 7.) As eccentricity increases, the profiles get narrower, and the peaks drop below 100%. At 20° eccentricity, the peaks average only 78%.

Performance at a given letter position depends on the relative position within the trigram. We refer to the center letter of the trigram as ‘middle,’ the one nearest the midline as ‘inner,’ and the one farthest from the midline as ‘outer.’

Fig. 5 shows separate profiles, averaged across subjects, for *inner*, *middle*, and *outer* letters. The profiles are for 0° eccentricity (i.e. central vision) and a presentation time of 200 ms. Notice that the highest accuracy (broadest profile) is for the letters furthest out from the midline. This surprising result replicates an effect re-

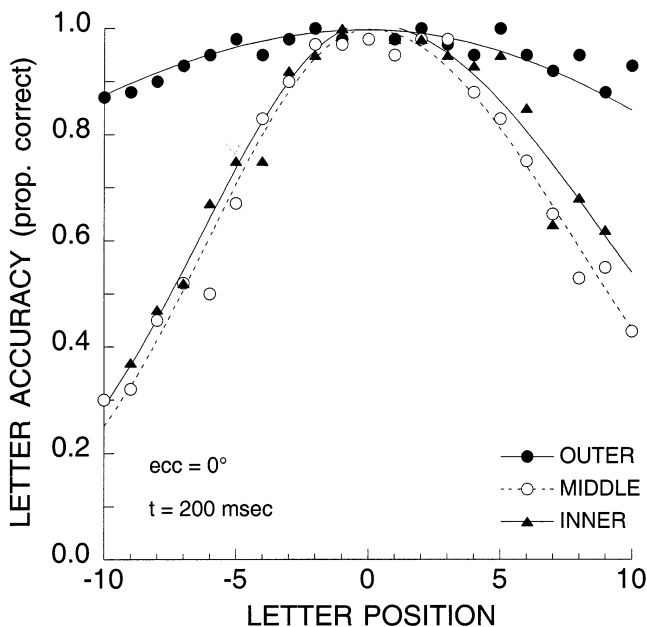


Fig. 5. Effect of letter position within the trigram: Separate visual-span profiles are shown for *inner* letters (the letter nearest the midline), *middle* letters, and *outer* letters (the letter farthest from the midline). (In Fig. 3, *n*, *s*, and *e* are inner, middle and outer letters, respectively for the sample trigram on the 5° line.) The profiles are for trigrams on the 0° horizontal line with 200-ms presentation times. Data have been averaged across the three subjects.

ported by Bouma (1970). The lowest accuracy (narrowest profile) is for the *middle* letters, although the difference between the profiles for *inner* and *middle* letters can be small, depending on presentation time and eccentricity.

Fig. 6 shows the effect of presentation time. Data, averaged across subjects, are shown for four of the eight presentation times. For the shortest exposure times (25 and 50 ms), peak accuracy drops below 100% and the profiles get narrower at all eccentricities.

Figs. 4–6 show that the shapes of the visual span profiles change with eccentricity, presentation time, and relative position within the trigram. Each profile can be summarized by the three parameters of the split Gaussian fits: amplitude, left standard deviation, and right standard deviation. Fig. 7 summarizes all our empirical visual-span profiles in this way. Each panel plots one of these three parameters as a function of exposure time for the four eccentricities. There are separate panels for each combination of the three Gaussian parameters and three relative positions within trigrams (inner, middle, and outer). The smooth curves fit to these summary values were used in the modeling described in Section 4.

Three details of the changing shape of the visual span can be observed in Fig. 7.

1. In central vision (0°), a peak amplitude of 100% is reached in 100 ms or less. In the periphery, peak performance takes longer to develop and falls short of 100%. For instance, at 20° eccentricity, peak performance on middle letters reaches only 69%. This latter observation implies that for someone reading at 20° in the lower visual field, visual information about letters flanked on both sides by other letters can, at best, support only 69% accuracy for the types of target letters studied here.
2. The breadth of the visual span grows more slowly and for longer exposure times than the amplitude. To the extent that people make use of the less reliable letter information in the tails of the visual span, increased exposure times beyond 100 ms are helpful.
3. The visual-span profiles tended to be slightly broader on the right than the left, as indicated by larger values of the right standard deviations in the Gaussian fits. For example, for 200-ms presentations, the ratio of standard deviations for profiles on the 0° line (averaged across subjects and relative trigram positions) was 1.30. Similar asymmetries for recognition of crowded letters have been reported by Bouma (1973) and Nazir et al. (1991).

In RSVP reading, words follow one another rapidly at the same retinal location potentially masking one another. The extent of this masking is likely to vary widely due to factors such as differences in word length and the degree of overlap of letter features. The trigram data we have described above are free of post-masking

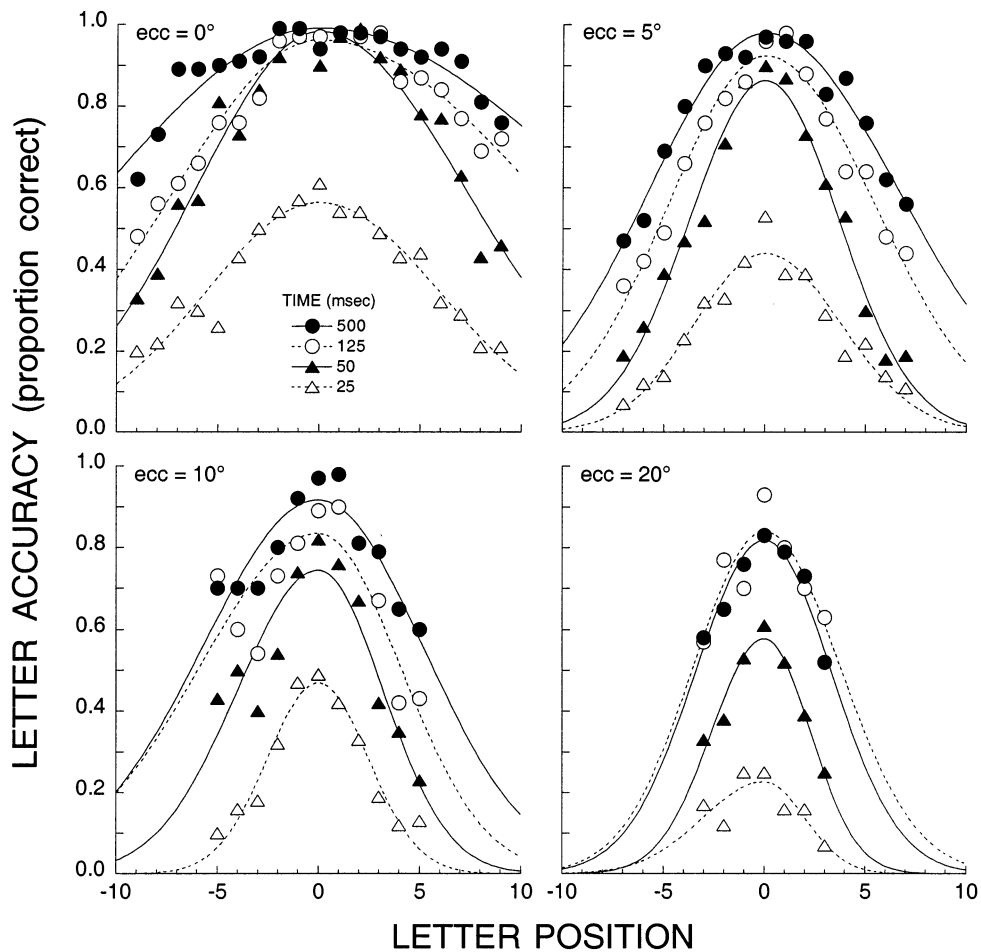


Fig. 6. Effect of trigram presentation time. Visual-span profiles are shown for four of the eight presentation times. Separate panels show profiles for the four eccentricities. Data have been averaged across the three subjects.

effects, and so represent an upper bound on the visual information for recognizing letters in reading. To estimate a lower bound due to masking, we performed limited measurements in which the trigrams were immediately followed by a mask composed of # # #.

Fig. 8 compares masked and unmasked visual span profiles for two exposure durations (40 and 200 ms) at 10° eccentricity. The effects of masking are pronounced for the short duration, but not for the longer one. Masking depresses the peak of the profile, but has relatively little effect on its breadth. The effects of masking are most marked at the shortest exposure times (< 80 ms) and are negligible for times longer than 125 ms.

### 3.3. Discussion

To put a number on the size of the visual span, we specify the retinal eccentricity and presentation time. Then, we can take the width of the appropriate spatial profile at a specified accuracy criterion (proportion-correct level).

Fig. 9 shows visual-span sizes computed in this way versus presentation time for different accuracy criteria (smaller spans for higher criteria). There are separate panels for the four eccentricities.

The visual spans get smaller in peripheral vision, and they get larger with prolonged presentation time. For example, for a 300-ms exposure and a 90% criterion, the visual span is about nine characters in central vision ( $\text{ecc} = 0^\circ$ ). This is close to the estimate from Legge et al. (1997a). But notice that if we insist on a 90% accuracy criterion, the visual span at 20° eccentricity is zero.

There are three limitations in expressing the size of the visual span in this way. First, the quoted size depends on the accuracy criterion, and the criterion is arbitrary. Second, the shape of the profile depends on whether we are considering inner, middle, or outer letters; but real text consists of a mix of these types<sup>8</sup>. Third, citing a spatial width does not reflect the dynamic character of the visual span.

<sup>8</sup> Here, we are considering the middle letters of trigrams to be representative of interior letters of words in general, and the inner and outer letters of trigrams as representative of ending letters of words.

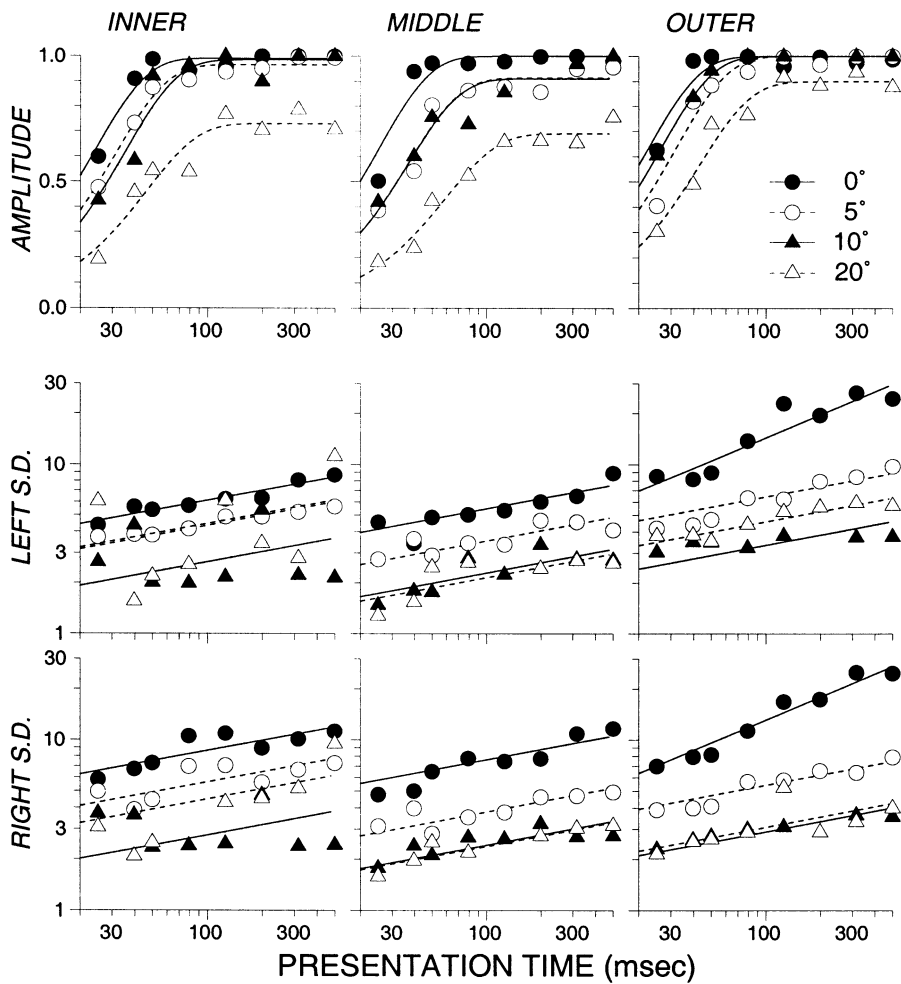


Fig. 7. Summary of trigram data from Exp. 2. Each visual-span profile can be summarized by the three parameters of the split Gaussian fits: Amplitude, Left S.D., and Right S.D. Each panel plots one of these three parameters as a function of presentation time for the four eccentricities. There are separate panels for each combination of the three Gaussian parameters and three relative positions within trigrams (*inner*, *middle*, and *outer*). The smooth curves fit to these summary values were used in the modeling (see text).

Fig. 10 illustrates a criterion-free way of plotting the rate at which information is transmitted by the visual span for an RSVP sequence of words. Information per word is expressed in bits. This quantity was derived from our visual-span profiles in the following steps: (1) Information transmitted at a given letter position on a profile was computed from percent correct recognition and ranged from 0 bits (for chance accuracy of 3.8% correct) to 4.7 bits (for 100% accuracy)<sup>9</sup>. (2) For words of a given length, assumed to be centered on the midline, we used visual-span profiles for the appropriate conditions (eccentricity, exposure time, and relative position within the word) to identify percent correct (and hence bits transmitted) associated with each letter position. We summed

<sup>9</sup> Percent correct letter recognition was converted to bits of information using letter-confusion matrices measured by Beckmann (1998). We computed the mutual information associated with confusion matrices. A plot of mutual information versus percent-correct letter recognition was well fitted by a straight line and was used to derive the values in Fig. 10.

the information transmitted at each letter position to compute the total information transmitted for words of that length. (3) We repeated this calculation for words of different lengths, drawn from the text passages used in Chung et al. (1998), and computed the overall informa-

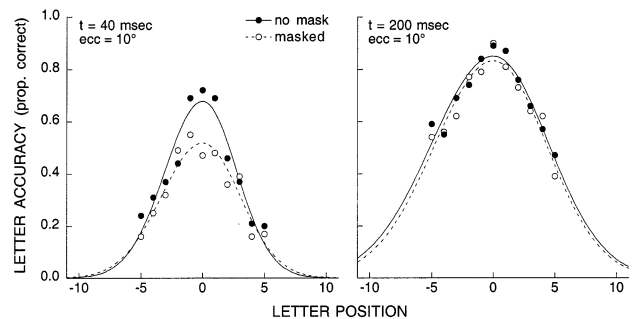


Fig. 8. Effect of masking. Visual-span profiles are shown for *no mask* and *mask* conditions for two presentation times, 40 ms (left) and 200 ms (right). The mask was composed of the characters # # # which immediately followed the trigram stimulus. The data are for 10° eccentricity and are pooled over two subjects (JSM and STC).

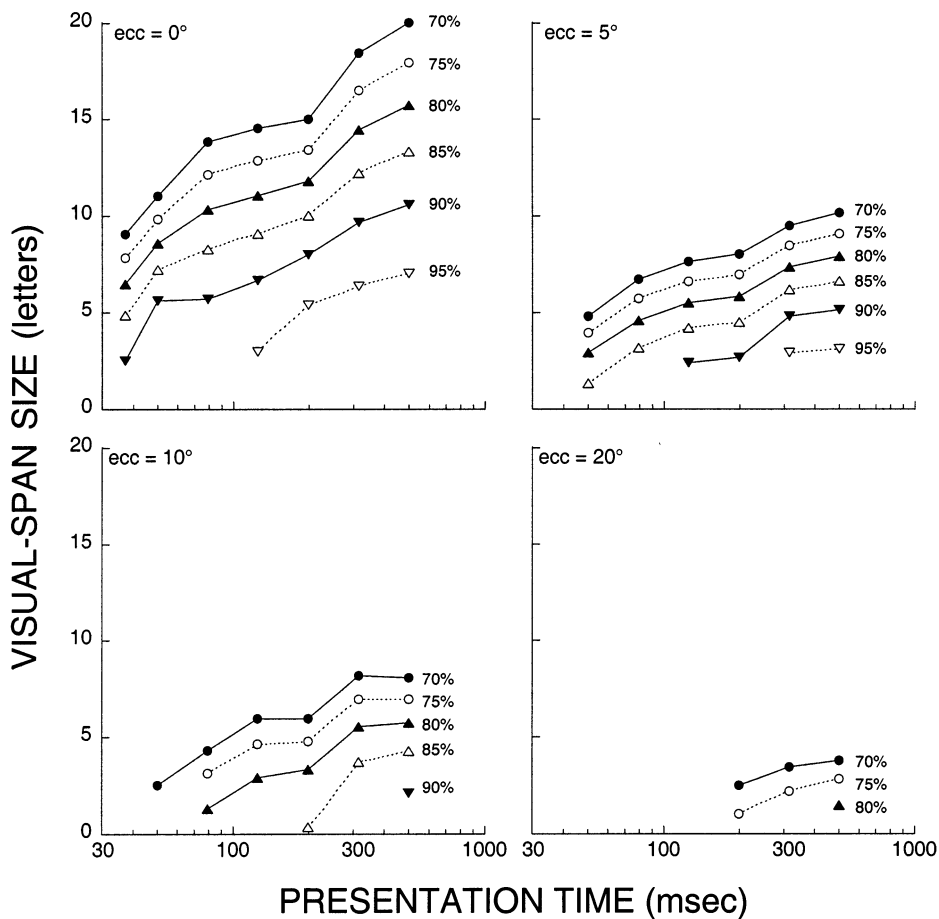


Fig. 9. The size of the visual span can be defined as the breadth of the visual-span profile for the appropriate eccentricity and presentation time. The numerical value will depend on the accuracy criterion adopted. This figure shows sizes for accuracy criteria ranging from 70 to 95% correct letter recognition. At 20° eccentricity, visual-span sizes for criteria at 85% and above were zero because the corresponding visual-span profiles peaked below 85% correct.

tion transmitted as a word-frequency weighted average across word length. These values are plotted against presentation time in Fig. 10.

Fig. 10 can be interpreted as showing the average amount of letter information per RSVP frame transmitted by the visual system. This is the bottom-up stimulus information available for word recognition or reading.

Fig. 10 also illustrates that information is transmitted faster in central vision. The curves roll over and approach their peak values at progressively longer exposures as we move into the periphery.

The spatial profiles measured in Exp. 2 show that the visual span varies with exposure time, retinal eccentricity, and even relative position within a text string. Clearly, there is no single visual span for a given subject, even at a given eccentricity.

Most of the variation of visual span with exposure time occurs for 100 ms or less. Consequently, this variation is not very important for eye-movement reading in which most fixations are much longer than 100 ms. But RSVP reading can sometimes employ exposure times of 100 ms or less, in which case the dynamics of the visual span

become relevant. Similarly, the effects of masking are largely confined to cases in which words follow one another within 100 ms.

The performance differences due to relative position within the trigram — inner, middle, outer — mean that the visual-span profile in reading will depend on the specific pattern of words and spaces encountered in a given fixation in text. The end letters of words (those adjacent to a space) will get a boost in recognition accuracy relative to interior letters. Clark and O'Regan (1999) exploited this property of recognition in their statistical analysis of word recognition. These fluctuations, due to local text properties, will average out across many text samples, yielding a characteristic accuracy level for a particular letter position.

Our key finding, however, is that the visual span, by whatever measure or index, gets smaller in peripheral vision. This finding confirms the conclusion of Exp. 1, using a very different method. Exp. 2 eliminated the likelihood of individual differences due to eye-movement or lexical-inference strategies. It refined characterization of the visual span by quantifying its spatial and temporal properties.

#### 4. Model: linking the visual span to RSVP reading speed

We now describe a model that takes the letter-recognition data from Exp. 2 as input and produces RSVP reading speeds as output. We will compare the model's performance in central and peripheral vision with the RSVP reading-speed data of Chung et al. (1998).

##### 4.1. Theory

The model reads words one-by-one as in RSVP. There are three stages (Fig. 11): (1) *fixation planning*: given the presentation time, the model plans the location and duration for each fixation, (2) *letter recognition*: visual-span profiles from Exp. 2 are used to model letter recognition performance for each fixation, and (3) *lexical matching*: the model matches the output from the letter-recognition stage with words in its lexicon.

##### 4.1.1. Fixation planning

The model knows the presentation duration for each word in the RSVP sequence. For purposes of modeling, we assume that the eyes must rest in one place for at least 200 ms before making a saccade, and that 50 ms are lost in executing the saccade. Fixations were constrained to last no longer than 500 ms.

According to these rules, if the presentation duration is less than 250 ms, only one fixation can be made, but the model can make one or more additional fixations if the duration is longer. For instance, for an RSVP

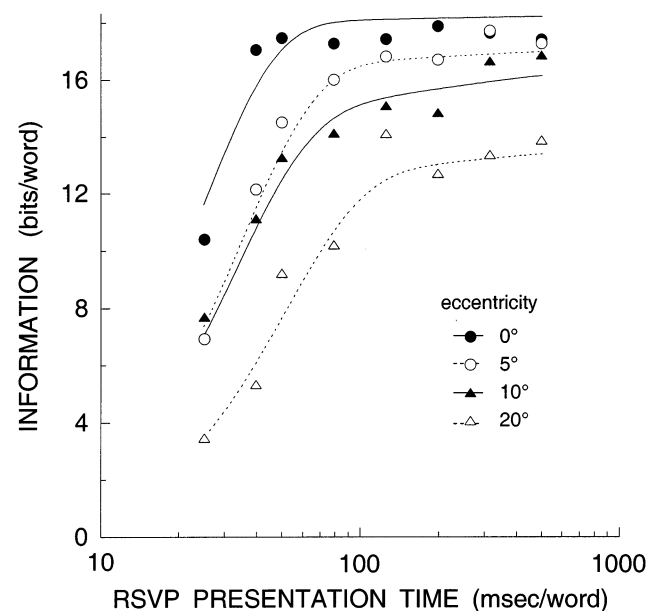


Fig. 10. Information transmitted by the visual span. A criterion-free method was used to compute the information per word transmitted by the visual span as a function of the presentation time per word in RSVP reading. See the text for an explanation.

exposure time of 500 ms, the model could make a 200-ms fixation, plus 50 ms for a saccade, followed by a second fixation of 250 ms. Alternatively, the model could extend the initial fixation to 225, 250, 275, ... up to 500 ms, reducing the time available for the subsequent fixations. The example in Fig. 11 has fixations of 250 and 200 ms.

We collected simulation data using two types of fixation strategy: *random* and *planned*. With *random fixations* the duration(s) and landing site(s) of each fixation were selected randomly, constrained so that the total time taken (including 50 ms for each saccade) equaled the presentation time.

With *planned fixations*, the duration and landing site for each fixation were selected to optimize the coverage of the word by the visual span according to the spatiotemporal properties of the visual span (measured in Exp. 2). The optimal sequence of eye movements was pre-computed for each eccentricity, presentation time, and word length, and was accessed via a look-up-table during the simulations.

##### 4.1.2. Letter recognition

To model letter-recognition, we first parameterized the spatiotemporal variations in the visual-span profiles from Exp. 2. Each visual-span profile was fit with an asymmetric Gaussian function according to

$$P(x) = \begin{cases} A \exp(-x^2/2\sigma_L^2) & \text{if } x < 0 \\ A \exp(-x^2/2\sigma_R^2) & \text{if } x \geq 0 \end{cases}$$

where  $P(x)$  is the probability of correct letter identification at letter position  $x$ ,  $A$  is the peak amplitude of the Gaussian, and  $\sigma_L$  and  $\sigma_R$  are the standard deviations (S.D.) of the Gaussian to the left and right of the peak.

Fig. 7 shows how  $A$ ,  $\sigma_L$  and  $\sigma_R$  vary as a function of time for the *inner*, *middle* and *outer* trigram positions at each of the four eccentricities tested. For modeling purposes, these data are fit with smooth functions as follows:

$$A = (1 - \lambda) \exp(t/\alpha)^{1.5}$$

$$\sigma_{L \text{ or } R} = at^b$$

The variation in amplitude,  $A$ , with time,  $t$ , was fit with a Weibull function with slope 1.5, with threshold,  $\alpha$ , and upper asymptote,  $\lambda$ , free to vary. The variation in standard deviation was well fit by a power law, with  $a$  free to vary and  $b$  set to 0.45 for outer letters at 0°, or 0.20 otherwise<sup>10</sup>. We can use these functions to interpo-

<sup>10</sup> For these latter cases, the value for  $b$  was determined by fitting the data from all the conditions together. For this reason some of the curve fits for the *inner* condition seem poor. However, data for the *inner* condition are not used in our current model, and thus the poor fits in these cases have no impact on our findings.

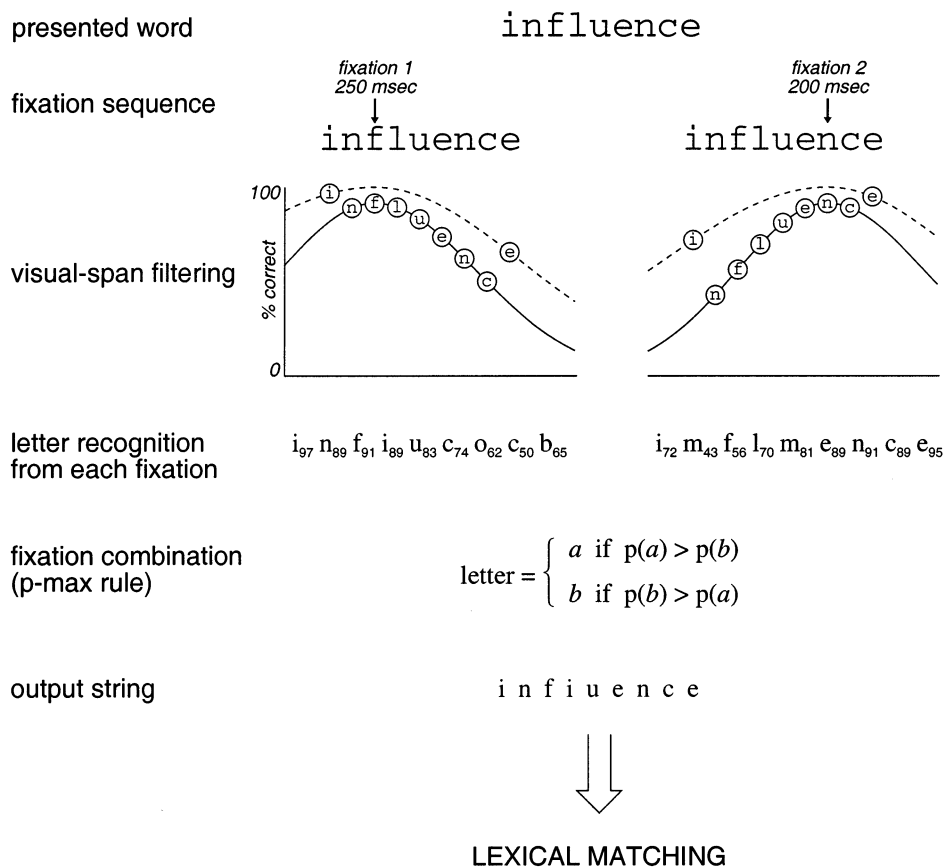


Fig. 11. Example of how the model encodes letter information from an input word. Here, the input word is 'influence' presented at 5° eccentricity for 500 ms. The model makes two fixations, lasting 250 and 200 ms, centered on *f* and *n*. (The intervening saccade is assumed to take 50 ms.) Letter-recognition accuracy is determined by appropriate visual-span profiles. The parameters of these profiles are taken from our data (summarized in Fig. 7). For the first fixation, we use a profile for a 250-ms presentation time and *middle* letters to determine recognition accuracy for the interior seven letters of the input word. Accuracy for the end letters is determined by a 250-ms profile for *outer* letters. In the diagram, the output letters from a fixation are subscripted with the accuracy of the corresponding slot on the visual-span profile. In the first fixation, for example, *c* is correctly transmitted through a slot with only 50% accuracy, but one of the *es* is incorrectly transmitted as *c* through a slot with 74% accuracy. Because transmission errors occur, the output strings from the two fixations are not identical. Discrepancies are resolved by the *p-max* rule: decide in favor of the slot with higher accuracy. This rule leads to a correct selection of *n* over *m* for the second letter of the word, but an incorrect selection of *i* over *l* for the fourth letter. The final output string 'influence' contains one transmission error. A lexical matching rule is then used (see text) to match this letter string to known words in the lexicon.

late within our data to generate the visual-span profile for any presentation time between 25 and 500 ms.

Within a single fixation, each letter of the stimulus word is 'filtered' through the appropriate visual-span profile. The profile specifies the probability of correct recognition of each of the letters in the word. The further out on the tails of the profile, the greater the chance of a letter-recognition error. Letter-recognition confusion matrices, collected in a separate study (Beckmann, 1998), were used to model human letter-recognition errors. We used 11 confusion matrices, which summarized letter confusions when overall recognition accuracy was 5, 15, 25, ..., 95, and 100% correct. For each letter in the word, we selected the confusion matrix with average %-correct closest to the %-correct for that letter position in the visual span. Then, the output letter identity was determined by choosing at random from the confusion-matrix row corresponding

to the presented letter. Visual-span profiles derived from *outer* trigram letters were used to model recognition performance for the first and last letters of the word. Profiles for *middle* trigram letters were used for the remaining letters.

On trials with two or more fixations, the model can generate different output strings for each fixation. We have considered two rules for combining the information across fixations to resolve these discrepancies: *ideal* and *p-max*. The ideal rule chooses the word  $w$  that maximizes  $p(w|w_1, w_2, \dots, w_n)$  where  $w_i$  is the letter string returned on the  $i$ th fixation. The *p-max* rule simply chooses the letter that came from the most reliable slot in the visual span. For example, in Fig. 11, the two fixations yield different output strings for the stimulus word *influence*. On the first fixation, the *l* of *influence* is filtered through a fairly reliable slot near the peak of the visual span. Nevertheless, an error is made,

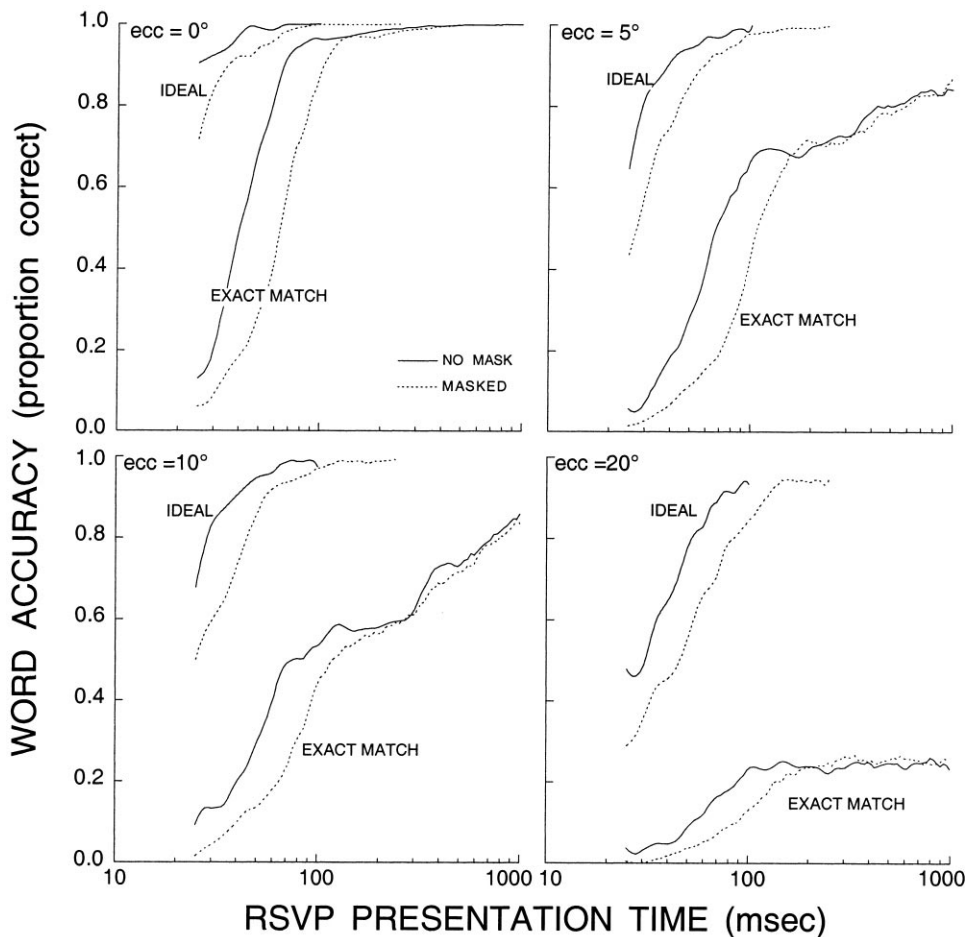


Fig. 12. The model's word accuracy (proportion correct) is plotted as a function of RSVP presentation time. Data are shown for simulations using the *ideal* and *exact match* lexical matching algorithms, based on letter-recognition data from Exp. 2 using masked and unmasked presentations.

and the output is  $i$ . On the second fixation,  $l$  is correctly reported. The  $p$ -max rule dictates that the model selects  $i$  because it came from a more reliable slot.

#### 4.1.3. Lexical matching

We have collected simulation data using two extreme rules for lexical matching: an *ideal rule* that makes optimum use of top-down information to identify the words even in the presence of letter recognition errors, and an *exact matching rule* that is unable to identify words in the presence of letter errors.

The *ideal rule* uses knowledge of the word frequencies in the lexicon and knowledge of the probabilities of letter confusions in implementing Bayes' theorem. Given the output string the ideal rule selects the word from the lexicon that maximizes:

$$p(\text{string}|\text{word})p(\text{word})/p(\text{string})$$

where  $p(\text{string}|\text{word})$  is calculated from the confusion matrices that were used to model the letter-recognition transmission errors, and  $p(\text{word})$  is the frequency of the word in the lexicon.  $p(\text{string})$  is the probability of the output string, and is determined by the frequencies of

the words in the lexicon, and the confusion matrices used for each letter position. In practice,  $p(\text{string})$  need not be evaluated to find the word that maximizes the expression<sup>11</sup>.

The *exact-matching rule* does not use any knowledge of word frequencies or letter confusions. This rule identifies words correctly only when there are *no* letter recognition errors. It requires an exact match between

<sup>11</sup> As an example of the ideal rule, suppose the output is the four-letter string *bont*. The ideal rule uses Bayes' theorem to decide on the most probable input word, given this output string. Suppose there are just five four-letter words:  $w_1$ : *boat*,  $w_2$ : *bond*,  $w_3$ : *sing*,  $w_4$ : *trip*,  $w_5$ : *very*. The ideal rule computes the probabilities of each of these words, given the output string *bont*. For instance for word  $w_j$ :

$$p(w_j|bont) = p(bont|w_j)p(w_j)/p(bont)$$

where  $p(bont) = \sum_{k=1, \dots, 5} p(bont|w_k)p(w_k)$ , and  $p(bont|w_j)$  is computed as the product of probabilities of each letter, based on the confusion matrices. For instance:

$$p(bont|boat) = p(b|b)p(o|o)p(n|a)p(t|t)$$

where  $p(n|a)$  is the probability of output  $n$  from the confusion matrix given input  $a$ .

the output letter string and a word in the lexicon. This simple rule can support rapid reading only when letter-recognition errors are rare.

#### 4.1.4. Simulating RSVP reading speed

Simulations were run to measure %-correct word recognition as a function of presentation time. Word presentation times ranged from 25 up to 1000 ms in 25-ms increments. Five hundred words were run for each presentation time. They were selected at random in proportion to their frequency of occurrence from the text corpus used by Chung et al. (1998). Plots of %-correct versus presentation time are shown in Fig. 12.

#### 4.2. Results

For the cases relevant to this paper, the effect on RSVP reading speed of random versus planned fixations is small<sup>12</sup>, so we will restrict attention to the

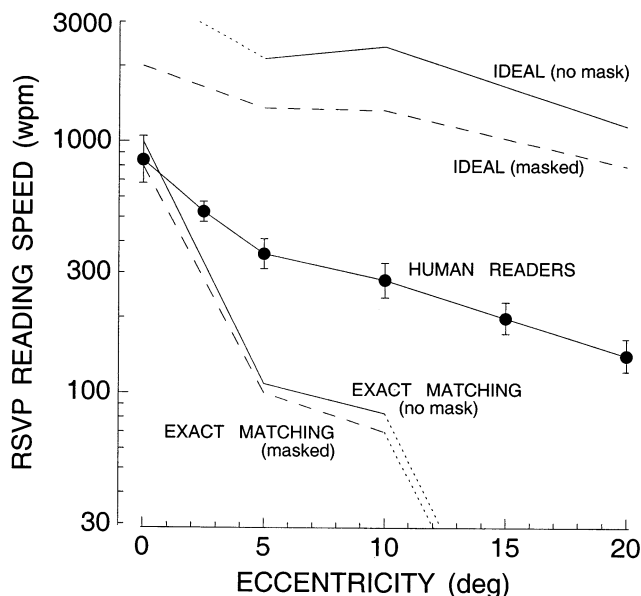


Fig. 13. Comparing predictions of the model to human reading speed. Mean RSVP reading speeds for six subjects are replotted from Chung et al. (1998) as a function of eccentricity. Predictions of two versions of the model are also shown. The Exact-Matching rule makes no use of lexical inference. The Ideal-Matching rule makes use of word frequencies and letter-confusion matrices to do lexical matching. For each of these rules, predictions are shown for masked and unmasked cases (refer to Exp. 2 and Fig. 8 for a description of masking conditions). For both the human and model data, reading speed was computed from the RSVP exposure times yielding 80% correct word recognition.

<sup>12</sup> For the ideal rule, planned fixations yielded reading speeds only 12% faster than random fixations (averaged across eccentricities and masking conditions). Since our ideal lexical-matching model could invariably achieve criterion accuracy in reading with a single fixation per word, this small difference indicates that position of the fixation within the word is not critical for good performance.

planned case only. Fig. 13 shows mean RSVP reading speeds versus eccentricity for six humans, replotted from Chung et al. (1998). The figure also shows the predictions of the model using either the *ideal* or the *exact* lexical matching rules. For each of these rules, predictions are shown for masked and unmasked cases (i.e. versions of the models in which the visual-span profiles were based on our masked or unmasked visual-span profiles in Exp. 2). For both the human and model data, reading speed was computed from the RSVP exposure times yielding 80%-correct word identification (by interpolation in Fig. 12). In computing the model's performance, no extra time was associated with the lexical-matching operations.

In central vision, the masked and unmasked versions of the model using the Exact-Matching rule predict reading speeds of about 800 and 1000 wpm, respectively, and tightly bracket the mean human value. This good fit in central vision means that the bottom-up visual data transmitted by the visual span, together with a very simple lexical matching operation can account for high-speed RSVP reading.

But the reading speed of the Exact-Matching model decreases more quickly than human performance in peripheral vision. At 20° eccentricity, neither the masked nor unmasked versions of the model is able to reach the 80% accuracy criterion for any RSVP exposure time. The reason for this poor performance in peripheral vision is that the visual-span profiles transmit letter information with low reliability, and it is relatively rare for all the letters of a word to be transmitted without errors.

In peripheral vision, humans perform better than the Exact-Matching model. Unlike the model, people apparently do some form of error correction on the fly. In other words they use some form of lexical inference to match letter strings containing errors to entries in the lexicon.

Is information available through lexical inference sufficient to account for human performance in peripheral vision? The answer is yes. This is shown by the reading speeds of the Ideal lexical matching model. This model puts an upper bound on the reading speed achievable through lexical inference. This ideal model easily outperforms humans. At 20° in peripheral vision, its reading speed is about equal to human performance in central vision.

The two versions of the model — one making ideal use of lexical inference, and the other doing no lexical inference at all — bracket human performance in peripheral vision. They leave open the challenge of discovering intermediate lexical-matching strategies that more nearly mimic human performance in peripheral vision.



### 4.3. Discussion

This model is similar to one proposed by Nazir et al. (1991) who were studying the optimal-viewing position effect in word recognition. But, to our knowledge, ours is the first model to forge a direct link between letter-recognition data and reading speed.

The model provides us with two important insights about reading. First, spatiotemporal limits on letter recognition, coupled with a simple lexical-matching operation, are sufficient to account for RSVP reading speed in central vision. Second, in peripheral vision, people must invoke some form of lexical inference because they achieve reading speeds unattainable with exact lexical matching.

Exps. 1 and 2 provided evidence for a quantitative difference between central and peripheral vision in the size of the visual-span profiles. Our modeling implies that there is also a qualitative difference; in peripheral vision, people use lexical inference to enhance reading speed, but in central vision they do not.

Why don't people use lexical inference in central vision as well? The performance of our ideal lexical-matching model shows that higher reading speeds are certainly possible in central vision. A likely explanation is that human lexical inference takes time. At 50–100 ms/word, typical of RSVP reading in central vision, there may not be enough time for lexical inference. In peripheral vision, the bottom-up processing takes longer because of the slower dynamics of the visual span and the more frequent need for multiple fixations. This extra time may permit use of slower lexical-inference processes. Alternatively, the need for error correction may force RSVP to slow down to accommodate the slower inference operations.

One obvious limitation of our model is that it does not apply to ordinary page reading. Although it does include within-word 'refixations', it does not include saccades from one word to the next. The RSVP formulation was convenient for us because it applied directly to the reading-speed data of Chung et al. (1998), and because the one-word-at-a-time analysis was easiest to implement.

Finally, note that the model presented in this paper differs from the Mr. Chips ideal-observer model (Legge et al., 1997b). Important differences include the following. (1) The present model uses data from human letter recognition as its visual-span profiles, while Mr. Chips uses visual spans with abrupt edges. (2) The present model takes time into account whereas Mr. Chips only indirectly involves time. As such, the present model provides an explicit prediction of reading speed in words per minute. (3) The present model applies to RSVP sequences of words, but Mr. Chips plans saccades through lines of text. (4) The present model takes into account letter-recognition errors and considers

how to deal with them in lexical matching. Mr. Chips assumes reliable letter recognition. (5) Different versions of the present model make varying use of lexical inference. By definition, Mr. Chips always does ideal lexical inference.

### 5. General discussion

This paper has adopted the theoretical view that letter recognition precedes word recognition in reading, and is fundamental to it. The visual span is a way of summarizing the relevant spatial and temporal properties of letter recognition. We have adopted this stance on the grounds of simplicity, recognizing that there is a long and unresolved debate about the perceptual units in reading (letters, words, or something else). Our findings indicate that a simple, bottom-up (letters first) model accounts for RSVP reading speed in central vision. Our evidence for a lexical-inference mechanism, operative in peripheral reading, may be symptomatic of additional processes that use word length and other clues to word identity.

Even if bottom-up, letter recognition forms the basis of high-speed reading, there appears to be another mechanism at work. This second mechanism relies on the standard layout of letters in text. Several studies have shown that unusual layouts adversely affect word recognition or reading speed. Examples include 'butterfly' text in which letter size increases in proportion to distance from fixation (Nazir, Jacobs, & O'Regan, 1998), text split across scotomas (Klitz, Legge, & Tjan, 2000; Klitz, 2000), and text with randomly oriented letters (Ortiz, Mansfield, & Legge, 1996). Nazir et al. (1998) have proposed that perceptual learning supports rapid, parallel identification of letters in standard text layout. But what is learned? One possibility is that a perceptual-learning mechanism associates letter reliabilities (values of expected percent correct) with letter positions at varying distances from the midline. These reliabilities are used explicitly in our model (the *p-max* rule) and could also play a role in fixation planning (Nazir et al., 1998).

We used two very different empirical methods for studying the visual span. The task in Exp. 1, involving sequential word recognition and multiple fixations within words, was more similar to regular reading than the letter-recognition task used in Exp. 2. In Exp. 1, we found that the reduction in visual span from central to peripheral vision paralleled changes in reading speed (Chung et al., 1998), but the results suffered from individual differences. These differences may have reflected alternative strategies used by our subjects for identifying words in peripheral vision. In Exp. 2, we used a method that precluded the use of likely strategies.

The data of Exp. 2 also provided us with a richer characterization of the visual span in terms of the dependence of letter recognition on retinal position, exposure time, and relative position within strings. To the extent that letter recognition is important in reading, these findings provide a fundamental data set.

We conclude by returning to the theoretical and clinical perspectives outlined in the Introduction. The theoretical question is how to link letter recognition to reading speed. From our data and model, we conclude that the key points are: (1) Retinal position, exposure time, and relative position within a string are key factors that limit letter-recognition accuracy. (2) The visual-span profile portrays the decline of letter-recognition accuracy outward from the midline. (3) Letter information conveyed by the visual span represents the visual information available for word recognition or reading. (4) For words that are long in relation to the breadth of the visual span, errors in letter recognition will necessitate multiple fixations and/or use of lexical inference. Both of these mechanisms for error correction will slow down reading. (5) When the error rate is low, typically true for simple texts in central vision, letter recognition forms the bottleneck for reading speed. (6) When the error rate is high, typically true in peripheral vision, letter recognition slows down reading speed, but some compensation is available through lexical inference.

From a clinical perspective, our results confirm the hypothesis that the visual span shrinks in peripheral vision. This shrinkage is an inescapable consequence of the spatial inhomogeneity of the visual field. For people with central scotomas, the reduction of the visual span narrows the letter-recognition bottleneck in reading. We conclude that the shrinkage of the visual span is a key factor in explaining the reading deficits of people with central-field loss.

## Acknowledgements

This research was supported by NIH Grant EY 02934 to Gordon E. Legge. We thank Tim Klitz and Bosco Tjan for useful discussion.

## References

- Beckmann, P. J. (1998). *Preneural factors limiting letter identification in central and peripheral vision*. Doctoral thesis, University of Minnesota.
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178.
- Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, *13*, 767–782.
- Beard, B. L., Levi, D. M., & Reich, L. N. (1995). Perceptual learning in parafoveal vision. *Vision Research*, *35*, 1679–1690.
- Bullimore, M. A., & Bailey, I. (1995). Reading and eye movements in age-related maculopathy. *Optometry and Vision Science*, *72*, 125–138.
- Clark, J. J., & O'Regan, J. K. (1999). Word ambiguity and the optimal viewing position in reading. *Vision Research*, *39*, 843–857.
- Chung, S. T. L., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research*, *38*, 2949–2962.
- Farah, M. J., & Wallace, M. A. (1991). Pure alexia as a visual impairment: A reconsideration. *Cognitive Neuropsychology*, *8*, 313–334.
- Faye, E. E. (1984). *Clinical Low Vision*. Boston: Little, Brown.
- Fine, E. M., & Rubin, G. S. (1999). Reading with central field loss: number of letters masked is more important than the size of the mask in degrees. *Vision Research*, *39*, 747–756.
- Foster, D. H., & Bischof, W. F. (1991). Thresholds from psychometric functions: superiority of bootstrap to incremental and probit variance estimators. *Psychological Bulletin*, *109*, 152–159.
- Kilgarriff, A. (1997). Putting frequencies in the dictionary. *International Journal of Lexicography*, *10*, 135–155.
- Klitz, T. S. (2000). *The effect of visual span size and simulated central scotomas on reading performance*. Doctoral thesis, University of Minnesota.
- Klitz, T. S., Legge, G. E., & Tjan, B. S., (2000). Saccade planning in reading with central scotomas: Comparison of human and ideal performance. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a perceptual process*. Amsterdam: Elsevier.
- Legge, G. E., Rubin, G. S., Pelli, D. G., & Schleske, M. M. (1985). Psychophysics of reading. II. Low vision. *Vision Research*, *25*, 253–266.
- Legge, G. E., Ross, R. A., Isenberg, L. M., & LaMay, J. M. (1992). Psychophysics of reading. XII. Clinical predictors of low-vision reading speed. *Investigative Ophthalmology and Visual Science*, *33*, 677–687.
- Legge, G. E., Ahn, S. J., Klitz, T. S., & Luebker, A. (1997a). Psychophysics of reading. XVI. The visual span in normal and low vision. *Vision Research*, *37*, 1999–2010.
- Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997b). Mr. Chips: An ideal-observer model of reading. *Psychological Review*, *104*, 524–553.
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics*, *16*, 578–586.
- Nazir, T. A., O'Regan, J. K., & Jacobs, A. M. (1991). On words and their letters. *Bulletin of the Psychonomic Society*, *29*, 171–174.
- Nazir, T. A., Jacobs, A., & O'Regan, J. K. (1998). Letter legibility and visual word recognition. *Memory and Cognition*, *26*, 810–821.
- O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler, *Eye movements and their role in visual and cognitive processes* (pp. 395–453). New York: Elsevier.
- O'Regan, J. K. (1991). Understanding visual search and reading using the concept of stimulus 'grain'. *IPO Annual Progress Report*, *26*, 96–108.
- O'Regan, J. K., Levy-Schoen, A., & Jacobs, A. M. (1983). The effect of visibility on eye-movement parameters in reading. *Perception and Psychophysics*, *34*, 457–464.
- Ortiz, A., Mansfield, J. S., & Legge, G. E. (1996). Reading text with randomly oriented letters. *Investigative Ophthalmology and Visual Science*, *37*, S167.
- Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science*, *206*, 468–469.
- Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research*, *16*, 829–837.
- Rubin, G. S., & Turano, K. (1994). Low vision reading with sequential word presentation. *Vision Research*, *34*, 1723–1733.

- Rumney, N. J., & Leat, S. J. (1994). Why do low vision patients still read slowly with a low vision aid? In A. Kooijman, et al., *Low vision — Research and new developments in rehabilitation* (pp. 269–274). Amsterdam: IOS Press.
- Trauzettel-Klosinski, S., Teschner, C., Tornow, R.-P., & Zrenner, E. (1994). Reading strategies in normal subjects and in patients with macular scotoma assessed by two new methods of registration. *Neuro-ophthalmology*, *14*, 15–30.
- Westheimer, G. (1999). Can peripheral visual acuity be improved by training? *Investigative Ophthalmology and Visual Science*, *40*, S586.
- Whittaker, S. G., & Lovie-Kitchin, J. (1993). Visual requirements for reading. *Optometry and Vision Science*, *70*, 54–65.
- Wilson, H., Levi, D., Maffei, L., Rovamo, J., & DeValois, R. (1990). The perception of form: Retina to striate cortex. In L. Spillmann, & J. Werner, *Visual perception: The neurophysiological foundations* (pp. 231–272). New York: Academic Press.