Effects of Temporal Modulation on Crowding Zone, Visual-Span Size, and Reading Speeds

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Caroline Haberthy, B.S.

Graduate Program in Vision Science

The Ohio State University

2015

Thesis Committee:

Deyue Yu, Ph.D., Advisor

Angela Brown, Ph.D.

Andrew Hartwick, O.D., Ph.D.

Abstract

Crowding, increased difficulty in identifying a target due to the proximity of adjacent objects, is the most important sensory factor determining the size of the visual span (the number of letters recognized without moving the eyes) leading to compromised performance in peripheral reading (Pelli et al., 2007; Whitney & Levi, 2011; He et al., 2013). The aim of the present study is to investigate the effect of temporal modulation on crowding reduction and reading speed improvement in peripheral vision. Crowding in peripheral vision can be reduced by temporally separating a target and flankers (Huckauf & Heller, 2004; Scolari et al., 2007; Greenwood et al., 2014). In the present study, we compared four temporal modulation patterns: 1) “moving window” (sequential exposing of component letters), 2) “moving scotoma” (sequential masking of component letters), 3) “flashing” (repeated exposing and masking of all letters simultaneously), and 4) “static” (no temporal changes). We investigated whether these temporal modulations can change the spatial extent of crowding, the size of visual span, and reading speeds in the periphery, and if so, which is the most beneficial. We found that compared to the static condition, the spatial extent of crowding was reduced in the moving window condition. None of the temporal modulations improved the size of the visual span and rapid serial visual presentation (RSVP) reading speed for print sizes larger than critical print size (CPS), the smallest print size that allows for maximum reading speed (MRS). However, both the moving window and moving scotoma conditions showed enhanced reading speed for print sizes smaller CPS (0.8^{o} and 1^{o}). The results suggested that temporal modulation is more effective for slow reading with print sizes closer to acuity threshold.

Acknowledgments

My sincerest thank you to Dr. Deyue Yu for her continuous guidance and assistance with my research and thesis. Additionally, I would like to thank Dr. Angela Brown and Dr. Andrew Hartwick for generously agreeing to participate in my thesis committee, for their time reviewing my research, and for their insightful questions and suggestions. This research was supported by NIH Grant T35-EY00715114.

Vita

June 2002 .... Northwest High School, Ohio

June 2004 .... Fontane Gymnasium Rangsdorf, Germany

June 2007 .... B.S. Business Administration, Shawnee State University

September 2007 to May 2008 .... Human Resources Administrative Assistance, Liebert Corporation

May 2008 to July 2009 .... Human Resources Coordinator, Liebert Corporation

May 2011 .... B.S. Biomedical Science, Shawnee State University

August 2011 to present .... Optometry Student, College of Optometry, The Ohio State University

Fields of Study

Major Field: Vision Science

Table of Contents

List of Figures

Introduction

Age-related macular degeneration (AMD), characterized by the progressive loss of central vision and relatively preserved peripheral vision, is one of the leading causes of blindness amongst older adults (e.g. Lim et al., 2012). Patients with AMD are forced to rely on their peripheral vision for everyday tasks. Reading, a critical task of daily living that relies heavily on central vision in normally-sighted people, becomes slow and difficult in patients with central vision loss (Legge et al., 1985; Legge et al., 1992; Fletcher et al., 1999). Given that reading difficulty is a major complaint of patients with central vision loss (e.g. Elliott et al., 1997), developing suitable reading rehabilitation to enhance reading speed is crucial for these patients.

It has been suggested that reading speed is primarily determined by fixation duration and saccade length (Parish & Legge, 1989). Patients with central vision loss do not necessarily spend longer duration on each fixation (Bullimore & Bailey, 1995, Crossland & Rubin, 2006) but the length of their saccades is consistently shorter leading to slower reading (Parish & Legge, 1989; Rumney & Leat, 1994; Bullimore & Bailey, 1995; Crosland & Rubin, 2006; Rubin & Feely, 2009). Saccade length has been tightly connected with the size of the visual span, the number of letters that can be recognized without moving the eyes (Legge et al., 1997; Legge et al., 2002). A large body of empirical data has confirmed that the size of the visual span imposes a sensory bottleneck on reading speed in normal and low vision (Legge, 2007; Legge et al., 2007).

There are three possible sensory factors determining the size of the visual span and reading speed: resolution (visual acuity), mislocation (position uncertainty), and crowding (Legge, 2007). Crowding (see Figure 1 for a demonstration) refers to the increased difficulty in identifying a target due to the proximity of adjacent objects (flankers). Among the three sensory components, crowding has been identified as especially problematic in peripheral vision (Flom et al. 1963; Bouma, 1970). It has been proposed that crowding, reflecting a failure of correct integration of target and flanker features, may be a result of pooling or averaging of information across an integration field (Parkes et al., 2001; Balas et al., 2009; Greenwood et al., 2009; Dakin et al., 2010; Freeman et al., 2012) or inappropriate feature substitution or migration due to positional uncertainty or unfocussed attention (Wolford, 1975; Chastain, 1982; Huckauf & Heller, 2002).

Figure 1. Demonstration of Crowding.

Both targets ("s") are equidistant from the fixation dot. When fixating the left dot, it is easy to discern the target "s" below. When fixating the right dot, it is more difficult to identify the same letter below because of the interference from flankers "o" and "u". This phenomenon is referred to as "crowding". In our study, stimuli were always presented at 10º below fixation.

Since crowding has been identified as the main sensory factor limiting the size of the visual span and since the visual-span size is tightly related to saccade length and reading speed, there has been a lot of interest in research investigating possible means of reducing crowding. In the clinical setting, text magnification is a simple and popular choice for aiding reading. However, text magnification can only assist with problems of reduced resolution in the periphery (Latham & Whitaker, 1996; Chung et al., 1998), and does not address the additional constraints imposed by crowding. With larger text, reading remains slower in the periphery compared to the fovea (Chung et al., 1998). In the laboratory, one intuitive notion to reduce crowding is the expansion of spacing between letters (Chung, 2002; Yu et al., 2007). However, despite the dependence of crowding on spacing, increasing letter spacing beyond the standard size does not lead to faster peripheral reading (Chung, 2002). As shown in Yu et al. (2007), the size of the visual span (the net sum of information transmitted considering all three sensory factors) does not increase for larger letter spacing, which explains the lack of corresponding improvement in reading speed.

Since crowding is also modulated by target-flanker similarity (Nazir, 1992; Kooi et al., 1994; Chung et al., 2001), alterations of the spatial feature resemblance between target and flanking letters have been studied as well. Despite the finding that crowding can be reduced when the target and flanking letters are of opposite contrast polarity, the mixed contrast polarity approach did not succeed in improving visual-span size or reading speed in the periphery (Chung & Mansfield, 2009).

In more recent years, researchers started to explore the effect of temporal manipulation on crowding. For example, motion such as forward zooming of one or more letters has been shown to reduce crowding, which is possibly due to better tagging of features to each individual letter, assisted by motion trajectory cues and/or potentially enhanced local attention (Husk & Yu, 2015). Varying the temporal separation (onset and offset) of target and flankers while all spatial properties remain unchanged has also been found to be effective in altering crowding. Greenwood et al. (2014) examined the effect of transient blink (selected elements blinked off and on in the middle of a presentation) on crowding, and found that crowding was reduced only when the target alone blinked, and remained the same when the flankers or all elements simultaneously blinked. They suggested that this asymmetric effect might be attributed to the separate activations of transient and sustained temporal channels within the visual system. The beneficial effect resulted mainly from transient target onset, which induced an additional stimulation of the transient channel to isolate the target signal. Huckauf & Heller (2004) conducted a study that systematically examined the effect of stimulus onset asynchrony between target and flankers on crowding. They found that crowding was alleviated (reflected by increased accuracy in target recognition) with larger temporal separation of target and flankers, especially for a flanker-preview condition (flanker onset prior to the target onset). The flanker-preview benefit was also confirmed by some later studies (Scolari et al., 2007; Greenwood et al., 2014).

In the present study, we aimed to assess whether temporal modulation can improve peripheral reading performance via a reduction in crowding. We examined four temporal modulation patterns (see Figure 2): “moving scotoma”, “moving window”, “flashing”, and “static”. The four temporal modulations differed with respect to the number of letters presented or masked at a given moment. In the moving scotoma condition, transient off-and-on manipulation was applied to letters one at a time from left to right. Likewise, in the moving window condition, transient on-and-off manipulation was applied to letters one at a time from left to right. In the flashing condition, transient on-and-off presentation was introduced to the entire letter array by repeatedly masking, then displaying all letters simultaneously. The static condition served as control and no temporal modulation was applied. Three measurements, the spatial extent of crowding (crowding zone), the size of the visual span, and rapid serial visual presentation (RSVP) reading speeds were obtained for each condition.

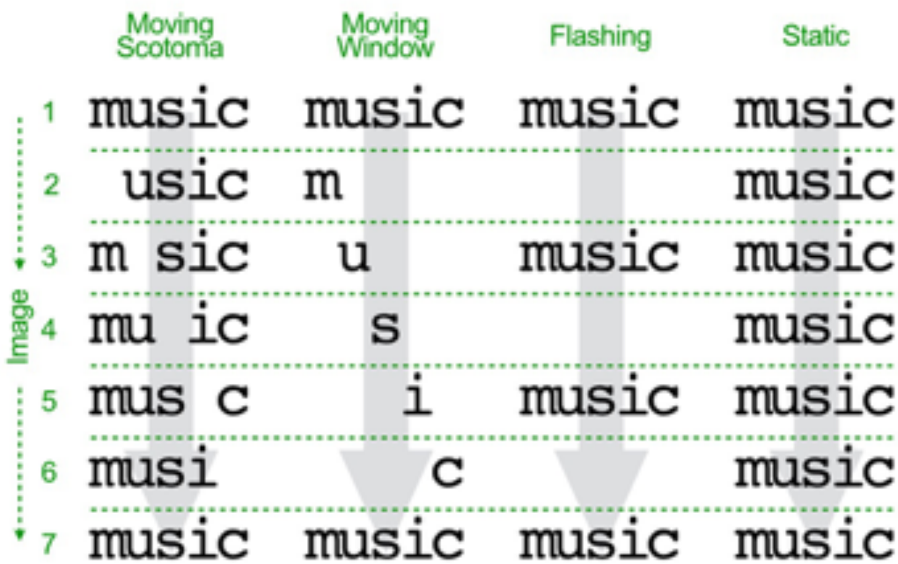


Figure 2. Samples of the Word "music" as Rendered in Four Temporal Modulation Conditions.

In our study, we used five-letter words to measure RSVP reading speed. Each temporal modulation was achieved by a seven-image sequence. We used trigrams, strings of three random letters, to measure the spatial extent of crowding and the size of the visual span (not shown in this figure). For trigram stimuli, each temporal modulation was achieved by a five-image sequence. Each column, illustrating the image sequence from top to bottom as time elapses, corresponds to one temporal modulation condition.

Based on the previous finding that transient target onset reduces crowding whereas simultaneous onset of target and flankers does not, we hypothesize that only the moving scotoma and moving window conditions may outperform the static condition on spatial extent of crowding. Since crowding has been identified to be the main sensory constraint on the size of the visual span and reading speed (Pelli et al., 2007; Whitney & Levi, 2011; He et al., 2013; Yu et al., 2014), we expected to find consistent changes in visual-span size and RSVP reading performance.

Methods: Participants

The study included six native English-speaking subjects between the ages of 21 and 25 years with normal or corrected-to-normal vision. None of the participants had any prior ocular surgeries or ocular conditions other than correctable refractive errors. None of the subjects reported difficulties reading. The tenets of the Declaration of Helsinki were followed throughout this study. The study protocol was approved by the Ohio State University Institutional Review Board. Each participant gave written informed consent prior to the data collection.

Methods: Apparatus and Stimuli

A ViewSonic Graphics Series G220fb CRT color monitor was used to present the stimuli. The frame rate and resolution were set to 85 Hz and 1280 × 1024 pixels (34.5 cm × 27.7 cm), respectively. The stimuli were generated using Matlab (version R2010a) and Psychtoolbox (Brainard, 1997; Kleiner et al., 2007). The Matlab program was administered via a MacBook Pro computer. Stimuli were displayed at 10º below fixation in Courier font, a fixed-width font. All stimuli were composed of lowercase English letters and were presented at -99% Weber contrast as black letters on white background (149 cd/m^2). Viewing was binocular at 40 cm viewing distance in a dark room.

Four temporal modulation conditions were applied to the stimuli: moving scotoma, moving window, flashing, and static condition (shown in Figure 2). Stimuli were displayed by five sequential image frames (for trigram stimuli) or seven sequential image frames (for five-letter word stimuli). For all conditions, the first and last image frames presented all letters. The rest of the image frames differed between the conditions. For the moving scotoma condition, as time elapses, a one-letter-wide invisible blind spot moves one letter position per image frame from left to right across the stimulus (i.e., masking one letter at a time). Similarly, for the moving window condition, a one-letter-wide window moves one letter position per image frame from left to right (i.e., showing one letter at a time). For the flashing condition, all letters are presented and masked simultaneously in an alternating manner. The static presentation serves as control and no temporal modulation was introduced (i.e., all letters are presented throughout all image frames).

Procedures and Data Analysis: Spatial Extent of Crowding

In the crowding task, subjects fixated a small dot in the center of the display while a trigram (random strings of three letters selected from the 26 lowercase English letters with replacement) was presented at 10º below fixation. Since the spatial extent of crowding is not affected by target size once it is above recognition level (Levi et al., 2002; Tripathy & Cavanagh, 2002; Pelli et al., 2004), a letter size of 1º (defined as x-height in lowercase) was used in the crowding task. Subjects were instructed to report the middle letter only. The presentation duration of each stimulus (sum of all five image frames) was 176 ms.

For each of the four temporal modulation conditions (moving scotoma, moving window, flashing, and static), accuracy of letter recognition was measured as a function of letter spacing. Six different letter spacings (0.8, 1.0, 1.4, 1.8, 2.5, 3.4 × x-width) were tested in this study. The standard spacing for Courier font is 1.16 × x-width. Subjects completed 40 trials per letter spacing and temporal modulation condition with a total of 960 trials. The data were fitted with a Weibull function. A criterion of 75% letter recognition accuracy was used to derive the spatial extent of crowding.

Procedures and Data Analysis: Size of Visual Span

Visual-span profiles were obtained with a letter-recognition task (e.g., Yu et al., 2007). Similar to the crowding task, the stimuli were trigrams presented with an exposure duration of 176 ms. A print size of 2.5º was selected to measure the visual-span size because this value exceeds the critical print size (CPS), the smallest print size that allows for maximum reading speed (MRS), for reading at 10° eccentricity (e.g., Chung et al., 1998). Subjects were asked to identify a trigram presented at different positions left and right of the midline at 10º eccentricity below fixation. The middle letter of a trigram was used to denote the trigram position. As shown in Figure 3, the position 0 marks the location directly below the fixation dot. There were a total of eleven trigram positions ranging from -5 to 5. Negative values refer to letter slots left to the midline; positive values refer to the positions right to the midline. Subjects were instructed to report all three letters from left to right. Letter identification was scored as correct only if the letter was reported at the correct location.

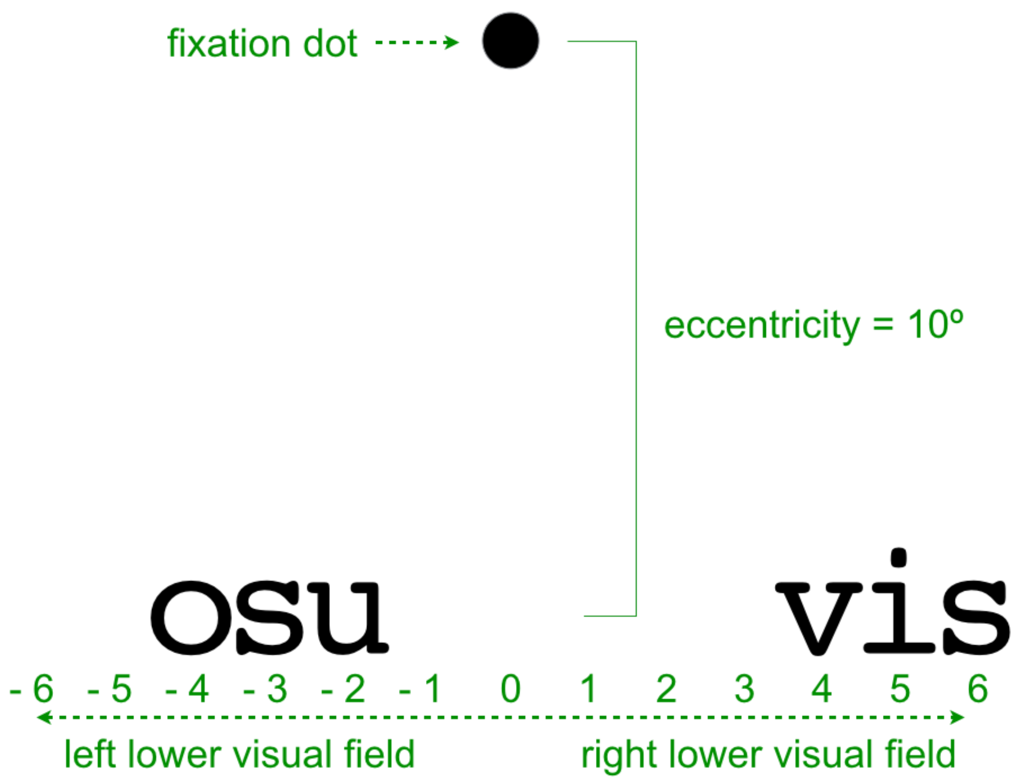


Figure 3. Examples of Trigrams Presented at Different Letter Positions.

Each trigram is presented at a letter position left or right of the midline at 10º below fixation. Trigram positions are defined by the position of the middle letter. For instance, the trigram “vis” is presented at position 5. The trigram “osu” is located at position -3 with the three letters positioned at -4 (“o”), -3 (“s”), and -2 (“u”), respectively. Position 0 corresponds to the location directly below the fixation on the midline.

Each subject completed a total of 880 trials (220 trials for each of the four temporal modulation conditions). Percent correct at each letter slot was accumulated from all letters (outer, middle, and inner letters of the trigrams in respect to fixation) presented at that location. Although we presented trigrams at eleven positions (from -5 to 5), only data from the central nine letter slots were analyzed because from positions -4 to 4, data were based on the trials from outer, middle, and inner letters of trigrams. As shown in

Figure 4, a plot of letter-recognition accuracy (proportion correct) as a function of letter position is referred to as a visual-span profile (Legge et al., 2001).

A split-Gaussian function was used to fit the visual-span profile with three parameters: the peak amplitude, the left-side standard deviation, and the right-side standard deviation (Legge et al., 2001). To quantify the size of the visual span, we converted proportion correct letter recognition into bits of information transmitted (Beckmann, 1998) and calculated the total amount of information transmitted by the nine slots of the visual-span profile (e.g. Legge et al., 2007; Yu et al., 2007). Information transmitted at any given letter position ranges from 0 bits corresponding to a chance accuracy of 3.8% to 4.7 bits for 100% accuracy.

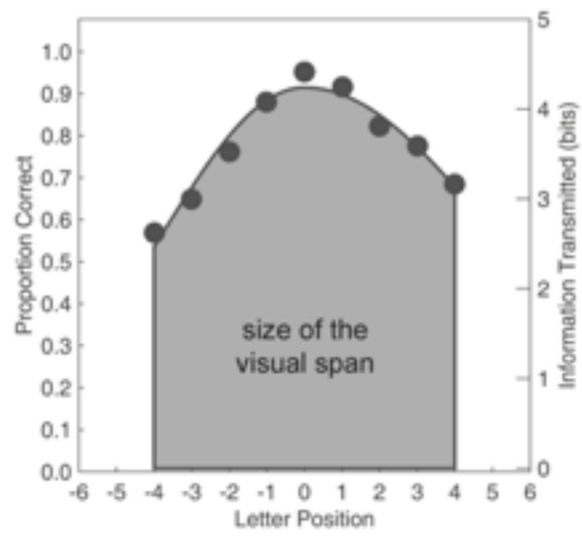


Figure 4. A Sample Visual-Span Profile.

A visual-span profile is a plot of letter-recognition accuracy (proportion correct) as a function of letter position accumulated across all trials under the same testing condition. The right vertical scale shows the conversion from the proportion of correct responses to information transmitted in bits.

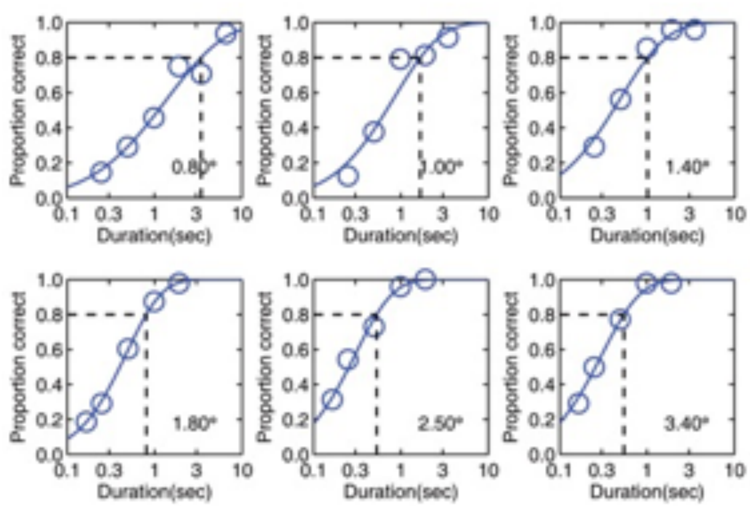
Procedures and Data Analysis: RSVP Reading Performance

Reading speed was measured using the rapid serial visual presentation (RSVP) paradigm (Rubin & Turano, 1992). In each trial, six five-letter words were randomly selected from a wordlist and presented one at a time in succession at 10º below a horizontal fixation bar (length is equivalent to the width of a five-letter word). We constructed the wordlist by compiling the most frequently used five-letter words in the English language at 8th grade reading level from multiple online sources. Plural nouns (i.e. "wools"), third-person conjugated verbs (i.e. "plays"), and conjugated verbs in past tense that are not used as adjectives (i.e. "lived") were excluded from the wordlist. The final list included 1,483 words. Prior to the start of the experiment, we asked each subject to review the wordlist and mark any unfamiliar words. None of the words were identified as unfamiliar by any of the subjects.

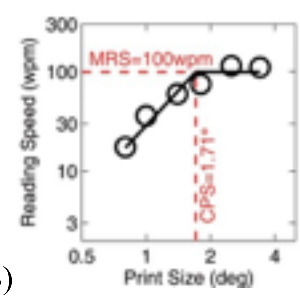
As demonstrated in Figure 2, presentation of each word consisted of seven image frames. During each trial, subjects were instructed to read the words aloud. Only horizontal eye movements along the fixation bar were permitted. We measured the proportion of words read correctly at five different exposure durations (increased in constant log steps) using the method of constant stimuli (eight trials per duration). For print sizes smaller than 1.5º, the exposure durations were 247, 494, 988, 1894, and 3459 ms/word. For print sizes larger than 1.5º, the exposure durations were 165, 247, 494, 988, and 1894 ms/word. When subjects' performance did not reach 80% accuracy at the longest exposure duration tested, additional eight trials were added with a longer exposure duration at the end of the last session.

To derive reading speed, we plotted the proportion of words read correctly as a function of presentation duration. We fitted the data with a Weibull function and used 80% as criterion to derive the threshold duration, from which we computed the reading speed in words per minute (see Figure 5A).

For each temporal modulation condition, reading speed was measured at six print sizes, 0.8º, 1.0º, 1.4º, 1.8º, 2.5º, and 3.4º. The total number of trials per subject was 960 plus a few additional trials at longer exposure durations for some subjects. As shown in Figure 5B, reading speed data were plotted as a function of print size. We fitted the data with a two-line function on log-log axes. The slopes of the two lines were constrained to 2.32 and zero, respectively (Chung et al., 1998). MRS and the CPS were given by the intersection of the two lines.



(A)



(B)

Figure 5. Example of RSVP Data Analysis.
continued on the bottom

Figure 5: continued
(A) The proportion correct of word recognition is plotted as a function of presentation duration for each print size. Threshold duration is determined using a criterion of 80% accuracy. (B) Reading speed computed from threshold duration is plotted as a function of print size. The data are fitted with a two-line function with one slope being 2.32 and the other being 0. The intersection of the two lines represents the CPS and MRS.

General Procedures and Data Analysis

Each subject took part in one practice session and four testing sessions. The trials obtained in the practice session were not included in our final data analysis. The order of the tasks (crowding, visual span, and RSVP reading) was randomized across testing sessions and subjects. For the crowding and the visual-span tasks, a random order of temporal modulation conditions (one block per condition per session) was constructed. The order of letter spacings (for the crowding task) or trigram positions (for the visual-span task) was randomized within each testing block. For the RSVP reading task, we first randomized the print sizes, then temporal modulation conditions, and finally exposure durations within each block. Subjects' eye movements during each testing trial were monitored by the experimenter. A trial was discarded and replaced when eye movements away from the fixation were detected.

We analyzed our data using repeated measures analysis of variance (ANOVA). The within-subject factor was the temporal modulation condition. Post-hoc tests were conducted when needed.

Results: Spatial Extent of Crowding

As shown in Figure 6, there was a significant effect of temporal modulation on the spatial extent of crowding ($F(3,15) = 7.832$, $p = 0.002$). The static condition yielded a spatial extent of crowding of 1.65±0.12 (SE) × x-width at 10º eccentricity below fixation. Among the four temporal modulation conditions, the moving window condition produced the smallest spatial extent of crowding (1.43±0.07 × x-width; $p < 0.05$). Compared to the static condition, the spatial extent of crowding was not significantly different for the flashing (1.57±0.10 × x-width) and the moving scotoma condition (1.64±0.09 × x-width).

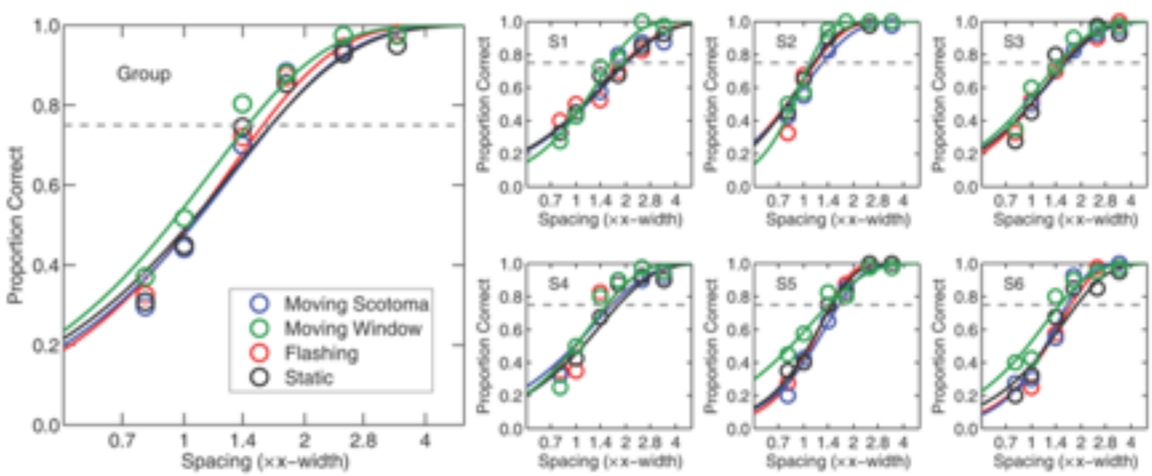


Figure 6. Results for Spatial Extent of Crowding for Four Temporal Modulation Conditions.

Proportion correct of letter recognition is plotted as a function of letter spacing (× x-width) for the group average (left) and for the six individual subjects (right). Data from each condition are fitted with a Weibull function. The dashed line represents the 75% letter recognition criterion to derive the spatial extent of crowding.

Results: Size of Visual Span

Visual-span profiles for the four temporal modulation conditions were plotted in Figure 7. The peaks of all the profiles occurred near letter position 0, and the right sides of the profiles were slightly broader than the left ($t(23) = 5.31$, $p < 0.0005$).

Our measure for the size of the visual span was the information in bits summed across the nine letter slots in the visual-span profile. Our statistical analysis showed that averaged across subjects, there was no significant effect of temporal modulation on the size of the visual span. In other words, the size of the visual span was the same across all temporal modulation condition at 10º eccentricity in the lower visual field. It was 32.52±0.31 bits for the moving scotoma condition, 32.28±1.20 bits for the moving window condition, 32.45±0.86 bits for the flashing condition, and 31.78±1.09 bits for the static condition, respectively.

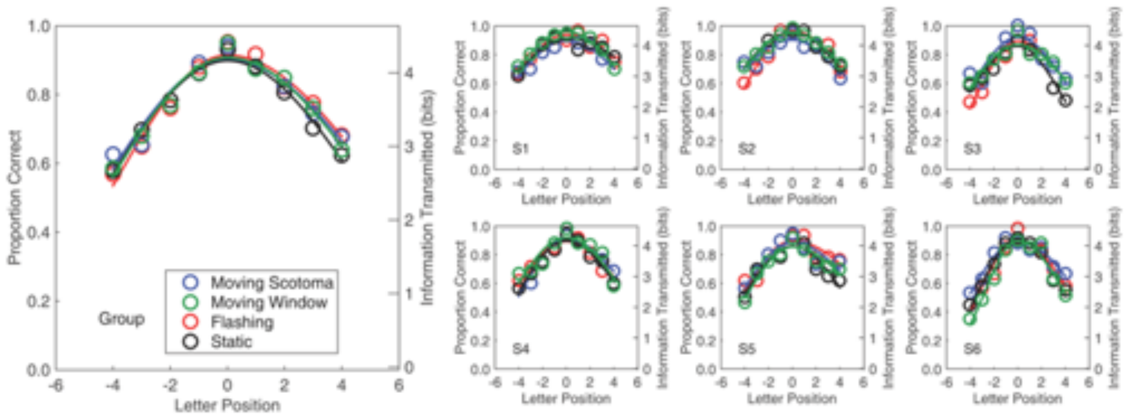


Figure 7. Results for Visual-Span Profiles for Four Temporal Modulation Conditions.

Proportion correct of letter recognition is plotted as a function of letter position for the group average (left) and for the six individual subjects (right). Each temporal modulation condition is represented by a different color. Data from each condition are fitted with split Gaussians.

Results: RSVP Reading Performance

Figure 8 shows RSVP reading speed in words per minute plotted as a function of print size for the four temporal modulation conditions. If 2.5° print size that was used to measure the visual-span profiles is above the CPSs for the four temporal modulation conditions, we would expect the MRS to have the same dependence on temporal modulation as the visual-span size measured at 2.5° print size. As shown in Figure 8, 2.5° is above the CPS across all testing conditions. Therefore, we expected the reading speed to remain at the same level across the four conditions at and above CPSs. However, The statistical analyses showed a significant effect of temporal modulation on both CPS ($F(3,15) = 5.81$, $p = 0.008$) and MRS ($F(3,15) = 14.36$, $p < 0.0005$). As shown in Figure 8, among all temporal modulation conditions, the moving window condition yielded the smallest CPS (1.65º±0.05º) and also the slowest MRS (102±12 wpm). For both CPS and MRS, no difference was found between the static condition (CPS = 2.04º±0.10º, MRS = 130±14 wpm) and the moving scotoma (CPS = 1.85º±0.10º, MRS = 127±12 wpm) and flashing (CPS = 1.95º±0.10º, MRS = 138±16 wpm) conditions.

Across all conditions and subjects, the CPS ranged between 2.29º and 1.46º. To examine the effects of temporal modulation on print sizes smaller than the CPS, we took a closer look at the two smallest print sizes, 0.8º and 1º. We found a significant effect of temporal modulation on reading speed ($F(3,15) = 5.49$, $p = 0.01$ for 0.8º and $F(3,15) = 6.19$, $p = 0.006$ for 1.0º). Post-hoc tests revealed that compared to the static condition

(16±5 wpm for 0.8º; 28±6 wpm for 1.0º), reading speeds were increased in both the moving window condition (28±4 wpm, $t(5) = 3.49$, $p = 0.018$ for 0.8º; 43±7 wpm, $t(5) = 3.76$, $p = 0.013$ for 1.0º) and the moving scotoma condition (27±7 wpm, $t(5) = 2.88$, $p = 0.035$ for 0.8º; 42±9 wpm, $t(5) = 3.77$, $p = 0.013$ for 1.0º). No significant improvement was found for the flashing condition for both 0.8º and 1.0º print sizes.

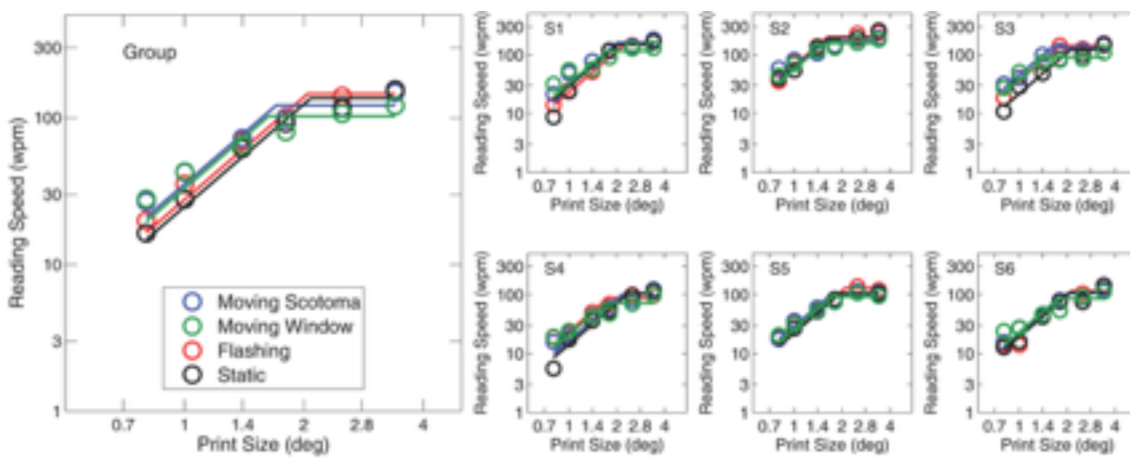


Figure 8. Results for RSVP Reading Speed for Four Temporal Modulation Conditions.

Reading speed in words per minute is plotted as a function of print size for the group average (left) and for the six individual subjects (right). Each temporal modulation condition is represented by a different color. Data from each condition are fitted with a two-line function on log-log axes. The slopes of the two lines are fixed to 2.32 and 0. The intersection of the linear functions marks the CPS and MRS for each temporal modulation condition.

General Conclusions

Temporal modulation has been shown to mitigate crowding for pattern recognition in the periphery (Huckauf & Heller, 2004; Scolari et al., 2007; Watson et al., 2012; Yu, 2012; Greenwood et al., 2014; Husk & Yu, 2015). The goal of our study was to investigate how temporal modulation (specifically varying the stimulus onset asynchrony among letters in a crowding setting) affects reading performance in peripheral vision. We compared four temporal modulation patterns: 1) "moving window" (sequential exposing of component letters), 2) "moving scotoma" (sequential masking of component letters), 3) "flashing" (repeated exposing and masking of all letters simultaneously), and 4) "static" (no temporal changes). We investigated whether these temporal modulations can change the spatial extent of crowding, the size of visual span, and reading speeds in the periphery, and if so, which is the most beneficial. We found that compared to the static condition, the spatial extent of crowding was reduced in the moving window condition. None of the temporal modulations improved the size of the visual span and reading speeds for print sizes larger than critical print size (CPS), the smallest print size that allows for maximum reading speed. However, both the moving window and moving scotoma conditions showed enhanced reading speed for print sizes smaller than CPS (0.8º and 1º). The results suggested that temporal modulation is more effective for slow reading, such as when print size is close to acuity threshold.

Discussion: Spatial Extent of Crowding

Consistent with observations from the previous studies (Huckauf & Heller, 2004; Scolari et al., 2007; Greenwood et al., 2014), we found that varying the stimulus onset asynchrony among peripherally presented letters can reduce crowding. Among the three non-static temporal modulation patterns that we investigated, only the moving window condition successfully reduced the spatial extent of crowding (by 13.3%) compared to the static condition. In this condition, only one letter was displayed for each of the three middle image frames. This spatial letter isolation likely explains the observed crowding reduction. However, even for the moving window condition, a substantial amount of crowding remained, which may be accounted for by the short exposure duration of each image frame (35 ms/image frame).

Short exposure duration may place a limitation on letter recognition through both within-frame processing and across-frame interference. It has been shown that recognition of isolated and crowded letters at a non-foveal retinal location reduces with decreasing exposure duration possibly following a non-linear (e.g. exponential) form for durations shorter than 100 ms (Bundesen & Harms, 1999; Legge et al., 2001). Huckauf and Heller (2004) investigated the effect of the stimulus onset asynchrony between target and flankers on crowding and found that target-flanker asynchrony may facilitate crowded target recognition in the periphery, depending on the amount of asynchrony. In the present study, the stimulus onset asynchrony between two adjacent letters was 35 ms.

According to the findings of Huckauf and Heller (2004), the benefit of 35 ms onset asynchrony (i.e. the reduction on the cross-frame interference) is likely small if any. This is indeed what we found.

Greenwood et al. (2014) examined the effect of transient blink on crowding. Transient blink can be considered as a form of target-flanker onset asynchrony. They found that crowding was reduced when transient blink was applied to the target alone and remained the same when the flankers or all elements simultaneously blinked. Consistent with their findings, we found no mitigation of crowding in the flashing condition in which the target and flankers were flashed on and off concurrently.

We also did not observe any significant reduction of crowding for the moving scotoma condition. Although this temporal modulation has not been tested previously in other studies, the finding of lack of crowding reduction was not surprising. We can identify three factors that likely influenced the effect of the temporal modulation on letter recognition: 1) spatial interference within each image frame, 2) exposure duration per image frame, and 3) interference across frames. The effect of temporal modulation may be accounted for by the sum of the three factors (one spatial and two temporal factors). For the spatial factor, there was reduced but non-zero spatial interference within each image frame as transient off-and-on manipulation was applied to letters one at a time from left to right in the moving scotoma condition. For the temporal factors, reducing exposure duration per image frame below a certain range deteriorates within-frame letter recognition through limited processing time and at the same time increases cross-frame interference on the letter recognition. In the moving scotoma condition, the exposure duration per image frame was short (35 ms/image frame), and therefore likely had a negative impact on letter recognition. Consequently, the effects of spatial and temporal factors canceled out and the resulting effect of the three factors for the moving scotoma condition did not lead to any reduction in crowding.

Discussion: Size of Visual Span

Our finding showed that temporal modulations did not increase the size of the visual span. This is curious because it has been shown that the visual-span size is primarily determined by crowding (e.g., Yu et al., 2014). Since the moving window condition reduced crowding by 13.3%, we expected to observe a similar beneficial effect in visual span. The differences in stimulus properties between the crowding and visual-span tasks included stimulus position and print size. In the crowding task, stimuli were always presented at 10º eccentricity directly below fixation (i.e., letter position 0). In the visual-span task, stimulus position ranged from -5 to 5. When only evaluating the performance at position 0, we still found no effect of temporal modulation on that part of the visual span.

Given that the standard letter spacing for Courier font is 1.16 × x-width, the standard spacing for 2.5º letter is 3.49º. If we express the spatial extent of crowding in the unit of degree, it ranges from 1.30º to 2.42º across subjects and conditions. The spacing of 3.49º used in the visual-span task greatly exceeded the spatial extent of crowding regardless of temporal modulation condition, which explained the lack of changes in the visual-span measurement at position 0. For the visual-span task, the average performance at position 0 was consistently close to ceiling (94%) across conditions. For the letter positions not tested in the crowding task, a similar argument may be true; or it may be that the spatial extent of crowding does not vary significantly with different temporal modulations.

Discussion: RSVP Reading Performance

According to the visual-span hypothesis, the size of the visual span imposes a sensory limitation on reading speed (Legge, 2007; Legge et al., 2007). Crowding is the major factor limiting the size of the visual span (Yu et al., 2014). Here, we found that temporal modulations did not increase the size of the visual span at a print size larger than the CPS. For print sizes smaller than the CPS, only the moving window condition successfully reduced the spatial extent of crowding. Based on these findings, we expected to find a similar pattern in RSVP reading speed at various print sizes. Although we did find a lack of improvement in reading performance at the print sizes near or larger than CPS and some beneficial effects of temporal modulation for reading performance at print sizes smaller than CPS, there were also some inconsistencies.

Firstly, although there was no concurrent change in the size of the visual span, the MRS was found to be slightly reduced for the moving window condition. The reduction in the MRS compromised the advantage of the associated reduction in CPS. The reduced CPS means that a smaller print size was needed for subjects to reach their MRS, which could be considered as a form of improvement. The discrepancy between the MRS and the size of the visual span was surprising. It is likely that subjects mainly used the central five letter slots of the visual-span profile when reading the five-letter words in the RSVP reading task. However, even when we examined the visual-span size for the central five letter slots only, there was still no effect of temporal modulation.

For print sizes below CPS (0.8º and 1º), two forms of temporal modulation (moving window and moving scotoma) outperformed the static condition on reading speed. The moving window presentation resulted in 61% and 127% faster reading speed for 1º and 0.8º print sizes, respectively. For the moving scotoma condition, the increase in reading speed was 53% for 1º print size and 97% for 0.8º print size. The presentation pattern of the moving window condition reduced the spatial extent of crowding which may have translated into improved reading performance. However, we were surprised to find that reading speed was increased in the moving scotoma condition given that crowding was not lessened under the same condition. This is the second discrepancy.

When considering only the spatial accounts (relationship of reading speed to the spatial extent of crowding and the size of the visual span), we cannot fully explain our findings. Given that we are investigating the effect of temporal modulation on various reading-related performance, it is reasonable to give the temporal factors a much closer look. In the letter recognition tasks, each trigram stimulus was presented for 176 ms and each of its image frames was presented for approximately 35 ms. In RSVP reading task, the duration per image frame (a total of seven image frames per word) varied depending on print size and performance level. For example, duration per image frame for the words presented at the MRS was 84 ms in the moving window condition and 66 ms in the static condition. To read with high accuracy at 1° print size, the duration per image frame needs to be about 200 ms for the moving window and scotoma conditions, and 309 ms for the static condition. When viewing crowded letters with very long exposure duration, letter recognition performance may be limited primarily by spatial interference such as crowding rather than processing in the temporal domain. Both moving window and moving scotoma conditions involve turning off letter(s) in the middle five image frames, which may reduce spatial crowding within each of the image frames. The effect only becomes salient when the exposure duration for each image frame is long enough (e.g. 200 ms). When duration per image frame decreases, the total amount of processing time for each letter decreases, and the temporal separation of image frames (i.e. image-onset asynchrony) becomes smaller (leading to increasing cross-frame interference). Below a certain exposure duration, temporal modulation such as the moving window condition can become harmful because it can reduce the processing time for each letter in the stimulus to a level not sufficient for letter recognition and interference across frames becomes more prominent. Therefore, a balance needs to be maintained to make a certain temporal modulation useful. When stimulus exposure is too brief, the additional temporal limitations introduced by a temporal manipulation may exceed its spatial benefit. When exposure duration is too long, the benefit on crowding reduction may be compromised by the slowness of the resulting reading speed.

Discussion: Clinical Implications

In our daily life, we encounter a diversity of print sizes from those below recognition level to those well above CPS. Small print sizes impose a particular difficulty for patients who suffer from central vision loss as visual acuity drops off quickly in the periphery. Clinically, we can enlarge small print to or above CPS electronically or optically with magnifiers. However, the enlarged text may run into the central scotoma or may be pushed further into the periphery where acuity deterioration, mislocation errors, and crowding increase. The method investigated in the present study may help enhance peripheral reading speed when enlarging print size is not an option. Despite the limited benefits revealed in the present study, temporal modulation may be a potential alternative in improving reading performance in brief, momentary reading tasks (e.g. news headlines, bank account balance, etc.) for people with central vision loss. Next steps in translating these findings into useful form of low-vision reading rehabilitation are to investigate questions such as whether the effect of temporal modulation is different for people with central vision loss, and whether the effect translates to page reading.

References

Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *J Vis, 9*(12), 13.11-18

Beckmann, P. J. (1998). Preneural factors limiting letter identification in central and peripheral vision. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(5241), 177-178.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis, 10*(4), 433-436.

Bullimore, M. A., & Bailey, I. L. (1995). Reading and eye movements in age-related maculopathy. *Optom Vis Sci, 72*(2), 125-138.

Bundesen, C., & Harms, L. (1999). Single-letter recognition as a function of exposure duration. Psychological Research, 62(4), 275-279.

Chastain, G. (1982). Confusability and interference between members of parafoveal letter pairs. *Percept Psychophys, 32*(6), 576-580.

Chung, S. T. (2002). The effect of letter spacing on reading speed in central and peripheral vision. *Invest Ophthalmol Vis Sci, 43*(4), 1270-1276.

Chung, S. T., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Res, 41*(14), 1833-1850.

Chung, S. T., & Mansfield, J. S. (2009). Contrast polarity differences reduce crowding but do not benefit reading performance in peripheral vision. *Vision Res, 49*(23), 2782-2789. doi: 10.1016/j.visres.2009.08.013

Chung, S. T., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Res, 38*(19), 2949-2962.

Crossland, M. D., & Rubin, G. S. (2006). Eye movements and reading in macular disease: further support for the shrinking perceptual span hypothesis. *Vision Res, 46*(4), 590-597. doi: 10.1016/j.visres.2005.05.015

Dakin, S. C., Cass, J., Greenwood, J. A., & Bex, P. J. (2010). Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *J Vis, 10*(10), 14. doi: 10.1167/10.10.14

Elliott, D. B., Trukolo-Ilic, M., Strong, J. G., Pace, R., Plotkin, A., & Bevers, P. (1997). Demographic characteristics of the vision-disabled elderly. *Invest Ophthalmol Vis Sci, 38*(12), 2566-2575.

Fletcher, D. C., Schuchard, R. A., & Watson, G. (1999). Relative locations of macular scotomas near the PRL: effect on low vision reading. *J Rehabil Res Dev, 36*(4), 356-364.

Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963). Visual resolution and contour interaction. *J Opt Soc Am, 53*, 1026-1032.

Freeman, J., Chakravarthi, R., & Pelli, D. G. (2012). Substitution and pooling in crowding. *Atten Percept Psychophys, 74*(2), 379-396. doi: 10.3758/s13414-011-0229-0

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proc Natl Acad Sci U S A, 106*(31), 13130-13135. doi: 10.1073/pnas.0901352106

Greenwood, J. A., Sayim, B., & Cavanagh, P. (2014). Crowding is reduced by onset transients in the target object (but not in the flankers). *J Vis, 14*(6), 2. doi:

10.1167/14.6.2

He, Y., Legge, G. E., & Yu, D. (2013). Sensory and cognitive influences on the training-related improvement of reading speed in peripheral vision. *J Vis, 13*(7), 14. doi: 10.1167/13.7.14

Huckauf, A., & Heller, D. (2002). What various kinds of errors tell us about lateral masking effects. *Visual Cognition, 9*(7). doi: 10.1080/13506280143000548A

Huckauf, A., & Heller, D. (2004). On the relations between crowding and visual masking. *Percept Psychophys, 66*(4), 584-595.

Husk, J. S., & Yu, D. (2015). The effect of motion on crowding: Zooming text. *J Vis, 15*(1). doi: 10.1167/15.1.17

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in psychtoolbox-3? (Vol. 36): Perception. ECVP Abstract.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spat Vis, 8*(2), 255-279.

Latham, K., & Whitaker, D. (1994). A comparison of word recognition and reading performance in foveal and peripheral vision. *Vision Res, 36*(17), 2665-2674.

Legge, G. E. (2007). *Psychophysics of Reading in Normal and Low Vision*. Mahwah, NJ & London: Lawrence Erlbaum Associates.

Legge, G. E., Ahn, S. J., Klitz, T. S., & Luebker, A. (1997). Psychophysics of reading--XVI. The visual span in normal and low vision. *Vision Res, 37*(14), 1999-2010.

Legge, G. E., Cheung, S. H., Yu, D., Chung, S. T., Lee, H. W., & Owens, D. P. (2007). The case for the visual span as a sensory bottleneck in reading. *J Vis, 7*(2), 9.1-15. doi: 10.1167/7.2.9

Legge, G. E., Hooven, T. A., Klitz, T. S., Stephen Mansfield, J. S., & Tjan, B. S. (2002). Mr. Chips 2002: new insights from an ideal-observer model of reading. *Vision Res, 42*(18), 2219-2234.

Legge, G. E., Mansfield, J. S., & Chung, S. T. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Res, 41*(6), 725-743.

Legge, G. E., Ross, J. A., Isenberg, L. M., & LaMay, J. M. (1992). Psychophysics of reading. Clinical predictors of low-vision reading speed. *Invest Ophthalmol Vis Sci, 33*(3), 677-687.

Legge, G. E., Rubin, G. S., Pelli, D. G., & Schleske, M. M. (1985). Psychophysics of reading--II. Low vision. *Vision Res, 25*(2), 253-265.

Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in amblyopic vision. *Vision Res, 42*(11), 1379-1394.

Lim, L. S., Mitchell, P., Seddon, J. M., Holz, F. G., & Wong, T. Y. (2012). Age-related macular degeneration. Lancet, 379(9827), 1728-1738. doi: 10.1016/S0140-6736(12)60282-7

Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Res, 32*(4), 771-777.

Parish, D. H., & Legge, G. E. (1989). *Psychophysics of reading: IX. The stability of eye position in normal and low vision.* Retrieved from http://legge.psych.umn.edu/pdf/read9.pdf.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nat Neurosci, 4*(7), 739-744. doi: 10.1038/89532

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J Vis, 4*(12), 1136-1169. doi: 10:1167/4.12.12

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. J Vis, 7(2), 20.21-36. doi: 10.1167/7.2.20

Rubin, G. S., & Feely, M. (2009). The Role of Eye Movements During Reading in Patients with Age-Related Macular Degeneration (AMD). *Neuro-Ophthalmology, 33*(3), 120-126. doi: 10.1080/01658100902998732

Rubin, G. S., & Turano, K. (1992). Reading without saccadic eye movements. *Vision Res, 32*(5), 895-902.

Rumney, N. J., & Leat, S. J. (1994). Why do low vision patients still read slowly with a low vision aid? In A. C. Kooijman, P. L. Looijestijn, J. A. Welling, & G. J. van der Wildt (Eds.), *Low vision: Research and new developments in rehabilitation* (pp. 269-274): IOS Press.

Scolari, M., Kohnen, A., Barton, B., & Awh, E. (2007). Spatial attention, preview, and popout: which factors influence critical spacing in crowded displays? *J Vis, 7*(2), 7.1-23. doi: 10.1167/7.2.7

Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Res, 42*, 2357-2369.

Watson, L. M., Strang, N. C., Scobie, F., Love, G. D., Seidel, D., & Manahilov, V. (2012). Image jitter enhances visual performance when spatial resolution is impaired. *Invest Ophthalmol Vis Sci, 53*(10), 6004-6010. doi: 10.1167/iovs.11-9157

Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn Sci, 15*(4), 160-168. doi: 10.1016/j.tics.2011.02.005

Wolford, G. (1975). Perturbation model for letter identification. *Psychol Rev, 82*(3), 184-199.

Yu, D. (2012). Jiggling the crowding away: improving letter recognition in peripheral vision. *J Vis, 12*(9), 2. ARVO Abstract. doi: 10.1167/12.9.330

Yu, D., Cheung, S. H., Legge, G. E., & Chung, S. T. (2007). Effect of letter spacing on visual span and reading speed. *J Vis, 7*(2), 2.1-10. doi: 10.1167/7.2.2

Yu, D., Legge, G. E., Wagoner, G., & Chung, S. T. (2014). Sensory factors limiting horizontal and vertical visual span for letter recognition. *J Vis, 14*(6), 3. doi: 10.1167/14.6.3