Title
Can reading-specific training stimuli improve the effect of perceptual learning on peripheral reading speed?

Authors
Bernard, JB
Arunkumar, A
Chung, STL

Publication Date
2012-08-01

DOI
10.1016/j.visres.2012.06.012

Peer reviewed

Vision Research 66 (2012) 17–25

Contents lists available at SciVerse ScienceDirect

Vision Research

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

Can reading-specific training stimuli improve the effect of perceptual learning on peripheral reading speed?

Jean-Baptiste Bernard *, Amit Arunkumar, Susana T.L. Chung

School of Optometry, University of California, Berkeley, Berkeley, CA 94720-2020, United States

A R T I C L E I N F O

Article history:
Received 10 December 2011
Received in revised form 21 May 2012
Available online 28 June 2012

Keywords:
Perceptual learning
Reading
Letter recognition

A B S T R A C T

In a previous study, Chung, Legge, and Cheung (2004) showed that training using repeated presentation of trigrams (sequences of three random letters) resulted in an increase in the size of the visual span (number of letters recognized in a glance) and reading speed in the normal periphery. In this study, we asked whether we could optimize the benefit of trigram training on reading speed by using trigrams more specific to the reading task (i.e., trigrams frequently used in the English language) and presenting them according to their frequencies of occurrence in normal English usage and observers' performance. Averaged across seven observers, our training paradigm (4 days of training) increased the size of the visual span by 6.44 bits, with an accompanied 63.6% increase in the maximum reading speed, compared with the values before training. However, these benefits were not statistically different from those of Chung, Legge, and Cheung (2004) using a random-trigram training paradigm. Our findings confirm the possibility of increasing the size of the visual span and reading speed in the normal periphery with perceptual learning, and suggest that the benefits of training on letter recognition and maximum reading speed may not be linked to the types of letter strings presented during training.

1. Introduction

People who lose their central vision due to retinal diseases such as age-related macular degeneration often have difficulty reading (Legge, 2007; Rubin, 2001). Indeed, reading is the primary goal for visual rehabilitation of the visually impaired (Elliott et al., 1997). The poor reading performance of people with central vision loss has been attributed to both oculomotor and early sensory factors (Legge, 2007). Early sensory factors include the limited visual acuity in the periphery (e.g., Mandelbaum & Sloan, 1947; Wertheim, 1980; Westheimer, 1979; Weymouth, 1958), crowding (e.g., Bouma, 1970; Levi, 2008; Pelli, Palomares, & Majaj, 2004) and slower processing time (Cheong et al., 2007). Legge et al. (2007) proposed that the *visual span*, the number of characters that can be recognized in a single fixation, represents the sensory bottleneck on reading. The evidence for the visual span as a primary limitation on reading speed was based on the significant correlations between the size of the visual span and reading speed for different print sizes, print contrast and testing eccentricities. These results further confirm the link between letter and word recognition (Legge, Mansfield, & Chung, 2001; Pelli, Farell, & Moore, 2003).

The correlation between the size of the visual span and reading speed implies that an increase in the size of the visual span should result in a corresponding increase in reading speed. Chung, Legge, and Cheung (2004) tested this prediction and showed that following four sessions of intensive training of trigram (random sequences of three letters) letter recognition at 10° eccentricity in the periphery in a group of normally sighted young adults, the size of the visual span became larger. More importantly, there was an accompanied improvement in reading speed at the same eccentricity, despite the fact that observers were not trained on a reading task. In that study, the perceptual learning task consisted of repetitive presentations of three random letters. However, English words do not comprise chunks of random letters. Instead, there are typical sequences of letters within words that are defined by strict linguistic rules, implying that there exist letter combination regularities in reading text. These regularities imply the existence of frequent letters (e.g., *t*, *e*), frequent bigrams (e.g., *on*, *th*) and frequent trigrams (e.g., *the*, *ion*) in reading materials, in contrary to low-frequency letters (e.g., *j*, *z*, *x*), non-existing bigrams (e.g., *bc*, *pq*) or non-existing trigrams (e.g., *bcx*, *aev*). It is suggested that observers' performance in word recognition could be linked to statistics of word elements, as shown by the effect of syllable frequency (Barber, Vergara, & Carreiras, 2004) and letter bigram frequency (Westbury & Buchanan, 2002) on word recognition. More generally, it has been shown that the visual system can learn the regularities of natural images and scenes through evolution and experience, which helps the visual system make use of its limited neural resources in an optimal way (Geisler, 2008).

* Corresponding author.
E-mail address: jb.bernard@berkeley.edu (J.-B. Bernard).

0042-6989/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2012.06.012

In the present study, we asked the question of whether we could optimize the benefit of expanding the size of the visual span following perceptual learning by replacing trigrams of random letters with chunks of three letters frequently encountered in reading. We hypothesized that by repeatedly presenting visual stimuli more specific to reading during perceptual learning, we could increase the effect of learning on the size of the visual span, which should then lead to higher reading speed. To test our hypothesis, we trained normally-sighted observers to recognize trigrams frequently encountered in the English language (trigrams that we called *reading trigrams*). To take into account reading statistics, the number of presentations of a trigram during training was directly linked to its probability of occurrence in the English language. We also made our learning procedure adaptive so as to increase the number of presentations for reading trigrams that were more difficult to recognize: during the procedure, the probability of presentation of a given trigram was modified based on whether or not the same trigram was correctly identified in a preceding trial. We expected that this specific training would be more effective in inducing an improvement in the size of the visual span, and thus lead to faster reading speed, compared with the conventional method of using trigrams of random letters.

To anticipate our results, our training method led to an increase in the size of the visual span, and was accompanied by an increase in reading speed. However, the magnitudes of improvement were not different from those of Chung, Legge, and Cheung (2004) in which trigrams of random letters were used for training. Our result suggests that the effect of the training was independent of the specific stimuli used.

2. Methods

2.1. Observers

Seven young adults (aged 18–23) with corrected-to-normal vision (20/20 or better acuity in each eye) participated in this study. None of the observers had prior experience in the tasks used in this study, or had participated in other experiments involving testing of peripheral vision. Refractive errors were corrected by glasses or contact lenses if necessary. Written informed consent was obtained from each observer after the procedures of the experiment were explained, and before the commencement of data collection. The experimental protocol was approved by the Institutional Review Board at the University of California, Berkeley.

2.2. Experimental design

The basic experimental design and training schedule followed closely those of Chung, Legge, and Cheung (2004). Each observer received a pre-test and a post-test with an intervening training procedure consisting of four sessions, scheduled on four consecutive days. The pre-test consisted of measurements of reading speed as a function of print size, followed by a measurement of the visual-span profile. The post-test consisted of a measurement of the visual-span profile, followed by measurements of reading speed as a function of print size. The pre- and post-test each lasted approximately 1.5–2 h. Training consisted of repeated measurements of the visual span profile, using the reading trigram stimuli (see below). Most training sessions were completed in an hour. All testings were performed binocularly at 10° below fixation (in the inferior visual field), with observers seated at 32 cm from the display.

Stimuli (trigrams and words) were generated on a Macintosh G4 computer with software written in Matlab 5.2.2 (The MathWorks, MA), using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997), and were presented on a Sony color graphics display monitor (Model# GDM-17E21, refresh rate = 75 Hz). All stimuli were rendered in Courier font as black letters (0.2 cd/m^2) on a white background (45 cd/m^2).

2.3. Reading speed measurement

As part of the pre- and post-tests, oral reading speeds were measured for six print sizes ranging from 0.7° to 4°, using the rapid serial visual presentation (RSVP) paradigm, i.e., words were presented one at a time in rapid succession, each for a fixed exposure duration (e.g., Chung, Mansfield, & Legge, 1998; Chung, Legge, & Cheung, 2004; Rubin & Turano, 1992, 1994). The psychophysical procedures and the sentence set used were identical to those used by Chung, Legge, and Cheung (2004). In brief, on each trial, a single sentence was chosen randomly, without replacement, from a pool of 2630 sentences extracted from classic literature. The observer, while fixating a fixation line, read aloud words presented one at a time at 10° below the fixation line. We used the Method of Constant Stimuli to present words at six exposure durations that spanned a range of approximately one log unit. An experimenter counted the number of words read correctly. There was no time pressure on the response and observers were free to complete verbalizing the words after the sentence was presented. The experimenter also monitored the eye movements of the observers during the presentation of words. A trial was discarded, and retested using a different sentence, when vertical eye movements away from fixation were detected. Across all observers, approximately 10% of the trials were discarded.

For each combination of print size and duration, we calculated the proportion of words read correctly by tallying across the six sentences presented for the same condition. Then for each print size tested, we used a cumulative-Gaussian function to fit the set of data relating the proportion of words read correctly with word exposure duration.[1] From the fitted function, we derived our criterion reading speed based on the word exposure duration that yielded 80% of words read correctly. Then, we plotted the criterion reading speed as a function of print size (see Section 3), and fit the dataset using a two-line fit (on log–log axes) to derive the two key measurements of reading performance – *maximum reading speed* (MRS) and *critical print size* (CPS, the smallest print size at which the maximum reading speed could still be attained). For the two-line fit, we followed the curve-fitting paradigm that was used in Chung, Legge, and Cheung (2004), with the slope of the first line constrained at 2.32 and that of the second one fixed at 0,[2] so that we could compare our results with those of Chung, Legge, and Cheung (2004).

2.4. Pre- and post-test visual span profile measurement

Visual-span profiles were measured using an identical letter-recognition task as in Chung, Legge, and Cheung (2004). In brief, on each trial, a trigram (a sequence of three lowercase letters randomly chosen from the 26 letters of the alphabet) was presented for 100 ms below the fixation target at 10° eccentricity (Fig. 1).

[1] Among a total of 84 sets of data (7 observers × 6 print sizes × pre/post-tests) that related the proportion of words read correctly with word exposure duration, 76 of them had the proportion correct of words read correctly ranging from an average of 0.25 to 0.89. For the other 8, the highest (raw) performance accuracy did not exceed 0.80, averaging 67% only. This occurred exclusively at the two smallest print sizes (0.7° and 1°). For these data-sets, the 80% criterion reading speed was extrapolated based on the cumulative-Gaussian function fit to the data.

[2] The log–log slope of the first line of the two-line fit was constrained at 2.32 as this was the value determined empirically in the study of Chung, Mansfield, and Legge (1998) in which they found that the slope of the first line did not vary systematically with eccentricity, and averaged 2.32 across all fitted data-sets (six eccentricities and six observers).

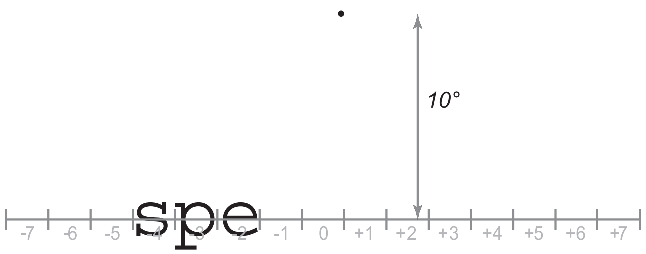

Fig. 1. A schematic cartoon illustrating the trigram letter-recognition task. The small black dot represents the fixation target. Trigrams were presented at 10° below the fixation target. In this example, the trigram "*spe*" is presented at a letter position (indexed by the middle letter) of −3 (three letter slots to the left of the vertical midline). The light gray horizontal lines and numbers indicating letter positions are for illustration purpose only, but were not presented on the actual display.

Observers verbally identified all three letters, from left to right. Eye movements of the observers were monitored to ensure that observers fixated at the fixation target. Letter size was 1.4× the critical print size, as determined from the pre-test reading task. We presented trigrams at 13 positions, indexed by the position of the middle letter, from six letter slots left of the vertical midline (letter slot 0 was 10° directly below fixation) to six letter slots right of the vertical midline. Each trigram position was tested 10 times in a random order within a block of trials, yielding a total of 130 trials tested in each block. A letter was scored as being identified correctly if and only if its order within the trigram was also correct. To calculate the overall accuracy of letter identification at each letter slot, we combined the identification scores across trials where the letter slot was occupied by the left, middle or right letter of a trigram. We then fit each set of data relating proportion-correct and letter position with a split-Gaussian function, representing the visual-span profile (Chung, Legge, & Cheung, 2004; Legge, Mansfield, & Chung, 2001). Curve-fitting was restricted to data within five letter slots left and right of fixation because the sixth letter slot left and right of fixation did not contain trials where the letter slot was occupied by the inner letter (the letter of a trigram closest to fixation). To quantify the size of the visual span, following Legge, Mansfield, and Chung (2001) and Chung, Legge, and Cheung (2004), we converted the identification accuracy at each letter slot to bits of mutual information transmitted by the visual span. According to Information Theory (Shannon, 1948), mutual information measures the amount of information that can be obtained about one random variable by observing another. In other words, it quantifies the dependence between the joint probabilities (the entropy or the uncertainty) of two events. With respect to our task of letter identification, the two events could be: what is the probability of an observer's response being an '*a*' given a stimulus letter '*o*'? Because there were 26 letters and there were many pairs of possible stimulus–response "joint events", the mutual information transmitted at a given letter slot ranged from zero bit for chance accuracy of 0.0384 to approximately 4.7 bits ($2^{4.7} = 26$) for perfect identification. To convert letter identification accuracy at a given letter slot to bits of mutual information transmitted, we used the following equation which was derived based on confusion matrices for single letter identification determined empirically (Beckmann, 1998):

$$\text{bits of information} = -0.037 + 4.676 \times \text{proportion correct of letter identification}$$

Then we summed up the total bits of information transmitted across all letter slots of the visual-span profile. This method of quantifying the visual span is akin to calculating the area under the curve.

2.5. Training

In Chung, Legge, and Cheung (2004), training consisted of 20 successive measurements of visual-span profile conducted over four sessions. In the current study, training was similar to that of Chung, Legge, and Cheung (2004), with the exception that letters within trigrams were not chosen at random. Instead, all trigrams presented were chunks of three successive letters of the 30,000 most frequently occurring English words. We shall refer to these trigrams with statistics that match those of English usage as *reading trigrams*. We used the frequency of English words as reported by the written part of the British National Corpus (*The British National Corpus, version 3 (2007)*, *distributed by Oxford University*) to calculate the frequency of each reading trigram in English literature. A total of 7239 reading trigrams were extracted, constituting approximately 40% of all possible *random* trigrams (letters chosen randomly). The most-frequently occurring reading trigram is "*the*", with a frequency greater than eight occurrences per 100 words.

In our study, we only presented the 2000 most-frequently occurring reading trigrams during training (for a list of the 2000 reading trigrams used for training in this study, please see the Supplemental information online).[3] Other reading trigrams have very low frequencies of occurrence and are very rare in the English language (e.g., the number 2001 trigram "*nuf*" has a frequency of less than 0.011 occurrence per 100 words). At the beginning of the training, we defined the initial probability of presentation of a given reading trigram for a specific letter slot (indexed by the middle letter of the trigram) as the frequency of that reading trigram divided by the sum of the frequencies of all the 2000 reading trigrams used during training. Subsequent probability of presentation varied for each letter slot during training according to the following rule: if the observer correctly identified all three letters of the reading trigram, the probability that the same trigram would be subsequently presented in the same slot was halved. However, when at least one letter error occurred, the probability of the same trigram being subsequently presented in the same slot was doubled. This adaptive procedure increased the number of presentation of reading trigrams that were difficult to recognize and reduced the number of presentation of readily identifiable reading trigrams. During training, observers were not aware that the presented trigrams were reading trigrams derived from words, and none of the observers noticed that only these reading trigrams, instead of random trigrams, were presented. As for the pre- and post-test visual-span measurements, eye movements of observers were monitored by the experimenter to ensure that observers' fixation was stable.

2.6. Data analyses

Each of the group-mean value reported in this paper represents the mean of the values across observers ±95% confidence intervals.

3. Results

Observers' performance for identifying letters in the reading trigrams during training is presented in Fig. 2, where performance accuracy (proportion-correct) for letter identification is plotted as a function of letter position. For these data, performance accuracy for each letter position was calculated for trials pooled across those in which the letter position was occupied by the left, middle or right letters of trigrams. Performance accuracy was only plotted for letter position (indexed by the middle letters of trigrams) with-

[3] The frequencies of occurrence of the 2000 reading trigrams used for training, selected based on the British National Corpus, highly correlate with the frequencies of occurrence of the same trigrams as appear in the set of sentences that we used for the reading speed measurements (correlation coefficient = 0.87).

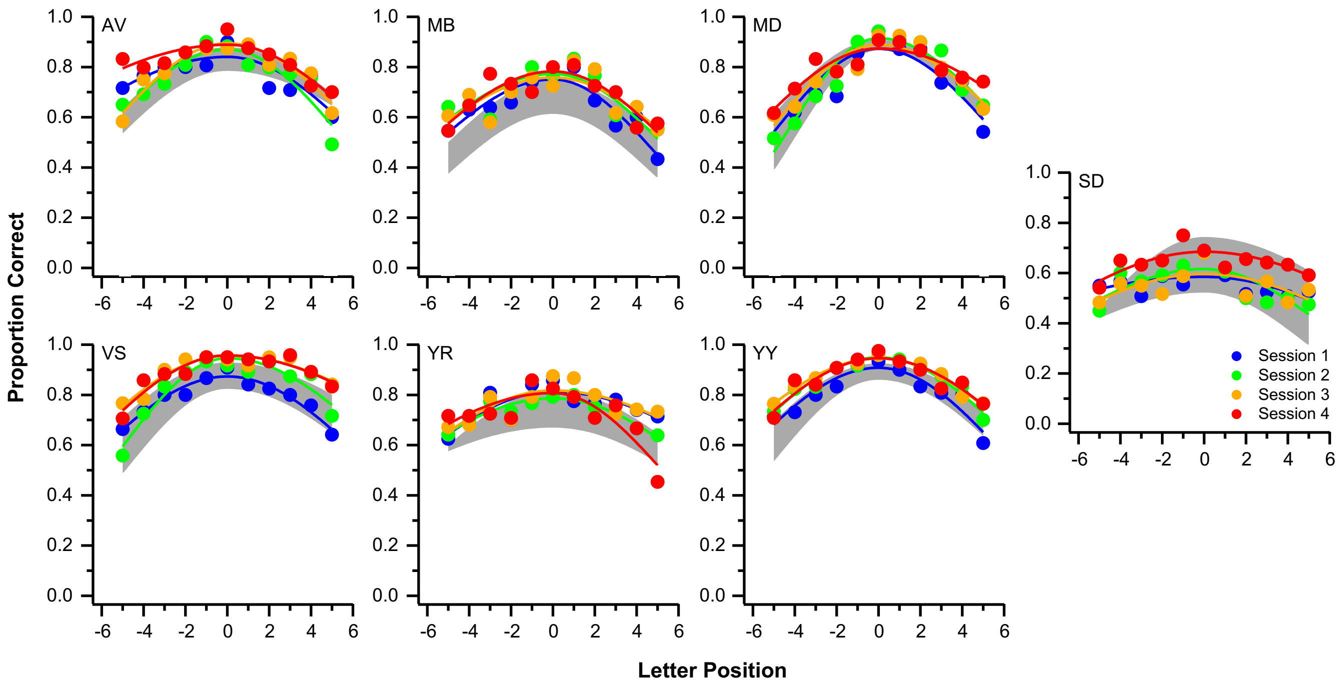

Fig. 2. Visual-span profiles, plots of proportion correct of letter-recognition as a function of letter position, obtained during training are shown for individual observers. Different colored symbols represent data obtained during different training sessions. The fitted curve through each set of data is the split-Gaussian function (see text for details). The shaded region in each panel represents the difference in performance accuracy between the pre- and post-tests.

in five letter slots right and left of the vertical midline, as letter slots outside this range did not contain the same number of presentations of letters. Each panel in Fig. 2 presents the data for one observer. The different colored symbols represent the overall performance for one training session, based on 1300 trials. The shaded region in each panel represents the difference in performance between the pre- and post-test measurement using *random* trigrams, with the lower bound of the shaded region representing the pre-test performance while the upper bound representing the post-test performance. To show more clearly the change in performance from pre-test to the training sessions, and to post-test, we plotted in Fig. 3 the difference in the size of the visual-span profile of a session relative to that of the pre-test, as a function of the testing session (pre-test, training and post-test). The small symbols represent data for individual observers while the larger symbols represent the group-average data. For the training data, there was an increase in the size of the visual span as training progressed (repeated measures ANOVA: $F_{(df=3,18)} = 5.63$, $p = 0.007$). Post-hoc pair-wise comparison using the Tukey HSD test showed that the only significant difference in the size of the visual span was between the first and the last training sessions. This result implies that the increase in the size of the visual span was progressive over time. Another interesting finding is that even though the pre- and post-tests were measured using random trigrams, not the reading trigrams as used during training, the post-test performance was significantly higher than that of the pre-test (average difference = 6.44 ± 1.53 bits, paired *t*-test: $t_{(df=6)} = 8.21$, $p = 0.0002$). This change in the size of the visual span was larger than that exhibited by the no-training control group (1.49 ± 1.25 bits, for the lower visual field measurements only) in Chung, Legge, and Cheung (2004) in which observers attended only the pre- and post-test sessions separated by four days and without any training. Further, there was a significant difference in the size of the visual span from the pre-test to the first training session (average difference = 4.51 ± 1.37 bits, paired *t*-test: $t_{(df=6)} = 6.44$, $p = 0.0007$);

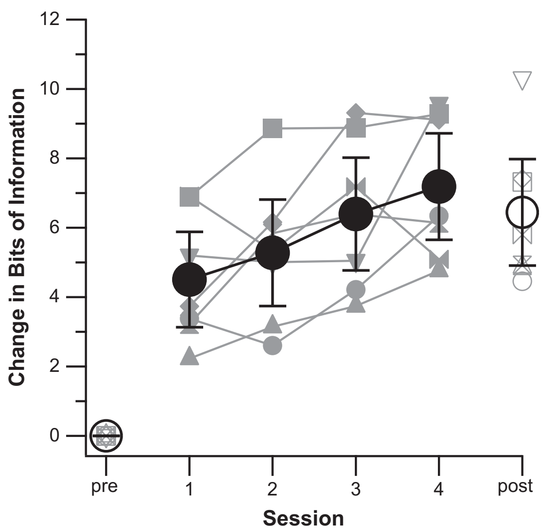

Fig. 3. The difference in the size of the visual span (quantified as bits of information) relative to that obtained at the pre-test, is plotted for the four training sessions (filled symbols) and the post-test (unfilled symbols). Small gray symbols represent data of individual observers. Large black symbols represent the group-averaged values. Error bars represent ±95% confidence intervals.

however, the size of the visual span between the last training session and the post-test was not significantly different (average difference = −0.75 ± 0.90 bits, paired *t*-test: $t_{(df=6)} = 1.62$, $p = 0.16$). Because the pre- and post-test, and the training sessions used different types of trigrams, the significance of these results will be discussed in Section 4 in relation to whether learning was specific to the (type of) trigrams used during training.

In Fig. 2, the lower and upper boundaries of the gray shaded region represent the pre- and post-test visual-span profiles of the

Table 1
Pre-test and post-test values for the size of visual span, maximum reading speed and critical print size for the seven observers.

Observer	Visual span size (bits)		Maximum reading speed (wpm)		Critical print size (°)	
	Pre	Post	Pre	Post	Pre	Post
AV	35.78	40.24	222.93	441.69	1.64	1.72
MB	25.70	33.02	109.74	194.54	1.48	1.73
MD	35.57	40.43	168.19	263.10	1.68	1.77
SD	23.02	33.34	249.65	432.98	1.62	1.71
VS	36.26	43.67	190.59	349.90	1.49	1.18
YR	31.58	37.39	134.98	192.60	1.17	1.43
YY	37.89	42.81	190.60	239.94	1.66	1.24

seven observers before and after training. These profiles were measured using random trigrams, following the pre- and post-test protocol of Chung, Legge, and Cheung (2004). For all observers, the post-test visual-span profile was above that of the pre-test, implying an improvement in letter-recognition performance across most letter positions following training. To facilitate the comparison of the visual-span profiles before and after training, we quantified the size of each visual span as bits of mutual information transmitted. Table 1 compares the size of the visual span for the seven observers before and after training. These data are also plotted in Fig. 5a where all the data points lie above the 1:1 equality line, implying an increase in the size of the visual span following training for all observers. Averaged across the seven observers, the size of the visual span increased from 32.26 ± 4.28 [95% confidence interval] bits (range: 23.02–37.89 bits) before training to 38.70 ± 3.17 bits (range: 33.02–43.67 bits) after training. In other words, the size of the visual span increased by an average of 6.44 bits (range: 4.46–10.31 bits) following training (paired *t*-test: $t_{(df=6)} = 8.21$, $p = 0.0002$).

Fig. 4 compares reading speeds for different print sizes before and after training. Most observers demonstrated higher reading speeds for a range of print sizes following training. To quantify the reading performance, we fit each set of reading speed vs. print size data using a two-line fit on log–log axes as described in Section 2.3. We were interested in comparing two parameters of the fitted function: the value on the y-axis that corresponded to the plateau, representing the MRS; and the intersection of the two lines, representing the CPS.[4] The pre- and post-test values of these two parameters for each observer are shown in Table 1. The comparison between the pre- and post-test measurements for MRS and CPS are also shown in Fig. 5b and c respectively. Averaged across observers, log maximum reading speed increased from 2.24 ± 0.09 (corresponding to 175.05 wpm, range: 109.74–249.65 wpm) at pre-tests to 2.46 ± 0.11 (286.46 wpm, range: 192.60–441.69 wpm) at post-tests, representing a 63.6% increase in maximum reading speed. Fig. 5b shows that all the data points comparing maximum reading speed before and after training lie above the 1:1 equality line, implying an improvement in the maximum reading speed following training for all observers (paired *t*-test: $t_{(df=6)} = 8.26$, $p = 0.0002$). However, there was no significant difference in the critical print size before (averaged = 1.54 ± 0.13°, range: 1.17–1.68°) and after (averaged = 1.54 ± 0.19°, range: 1.18–1.77°) training, as shown by the scattering of data points above and below the 1:1 equality line in Fig. 5c (paired *t*-test: $t_{(df=6)} = 0.04$, $p = 0.97$).

3.1. Comparison with the results of Chung, Legge, and Cheung (2004)

A central question of this study was whether training using reading trigrams extracted from the frequently used English words was more effective in enlarging the visual span and thus improving reading speed than using random trigrams. Given that the experimental procedures, analyses and parameters of interest followed closely those of Chung, Legge, and Cheung (2004), we were able to compare our results with those of Chung, Legge, and Cheung (2004). For the comparison reported here, we only included results from observers who were trained in the lower visual field (also at 10° eccentricity) in Chung, Legge, and Cheung (2004).

3.1.1. Size of the visual span

The magnitude of improvement following perceptual learning has been suggested to depend on the initial level of performance (Fahle & Henke-Fahle, 1996; Yu et al., 2010). For the size of the visual span, the pre-test value averaged 32.26 ± 4.28 bits in this study and 31.27 ± 3.91 bits in Chung, Legge, and Cheung (2004). These values were not different from each other (two-sample *t*-test: $t_{(df=11)} = 0.44$, $p = 0.67$). Following training, the increase in the size of the visual span, defined as the difference in bits between pre- and post-tests, averaged 6.44 ± 1.53 bits in this study, and was not statistically different from an average of 7.37 ± 0.82 bits in Chung et al. (two-sample *t*-test: $t_{(df=11)} = 0.99$, $p = 0.34$; see also Table 2).

3.1.2. Maximum reading speed

Just as for the size of the visual span, pre-test maximum reading speeds were similar between this study and Chung, Legge, and Cheung (2004). The pre-test log maximum reading speed averaged 2.24 ± 0.09 for this study and 2.28 ± 0.09 for Chung, Legge, and Cheung (2004) (two-sample *t*-test: $t_{(df=11)} = 0.62$, $p = 0.55$). Following training, the improvement in the maximum reading speed, defined as the ratio of post- to pre-test value, averaged 63.6 ± 18.6% in this study, and 45.7 ± 13.8% in Chung et al., which was not statistically different from each other (two-sample *t*-test: $t_{(df=11)} = 1.62$, $p = 0.13$; see also Table 2).

3.1.3. Critical print size

The average of the pre-test CPS values was 1.54 ± 0.13° for our observers, which was larger than the averaged value of 1.15 ± 0.13° in Chung, Legge, and Cheung (2004) (two-sample *t*-test: $t_{(df=11)} = 4.12$, $p = 0.002$). However, in both experiments, there was no significant change in CPS following training.

4. Discussion

In a previous study, Chung, Legge, and Cheung (2004) studied the effects of repeated presentations of trigrams composed of three random letters at 10° eccentricity in the periphery, on the size of the visual span and reading speed of six observers. This training method significantly improved the size of the visual span (averaged 7.37 ± 0.82 bits) and the maximum reading speed of all observers (averaged 45.7 ± 13.8%). In the present study, we hypothesized that the improvements on the size of the visual span and reading speed could be further enhanced by using reading trigrams that match the frequently occurring regularities of how letter combinations appear in English usage. Contrary to our prediction, we found that the improvements following training with

[4] We acknowledge that the two-line fit, with the log–log slope of the first line constrained at 2.32, did not seem to fit some data-sets adequately in Fig. 4. It is possible that other functions would have been more appropriate for our data, for example, if we did not constrain the slope of the first line, or if we used an exponential fit (Cheung et al., 2008). However, we adhered to the same fitting function (with the slope of the first line constrained at 2.32) as used by Chung, Legge, and Cheung (2004) so that we could compare our results with theirs. Different fitting methods would likely yield similar maximum reading speed, but the critical print size could be different (as the definition of the critical print size could vary for different fitting methods), which could make it difficult for us to interpret and compare the results between our study and that of Chung, Legge, and Cheung (2004).

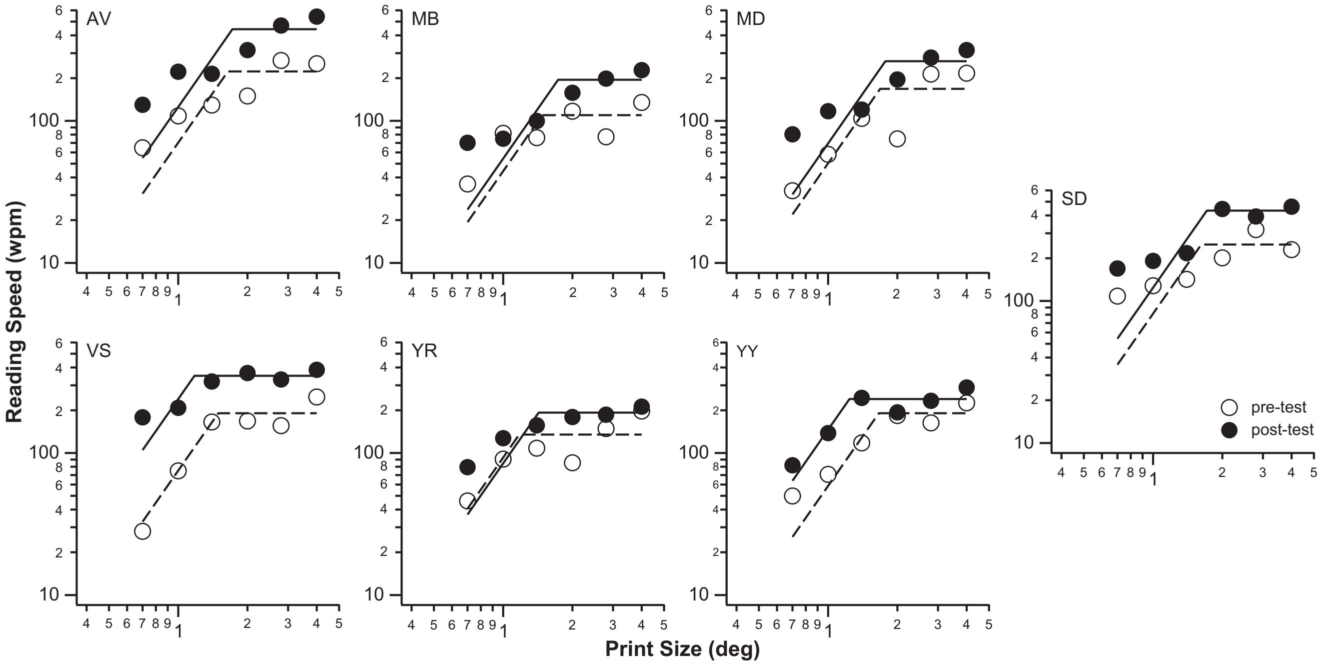

Fig. 4. RSVP reading speed (words per minute, wpm), is plotted as a function of print size (deg) for the seven observers. Unfilled symbols represent data obtained during the pre-test while filled symbols represent data obtained during the post-test. The straight lines through each set of data represent the two-line fit (see text for details) from which the maximum reading speed (MRS) and the critical print size (CPS) were derived.

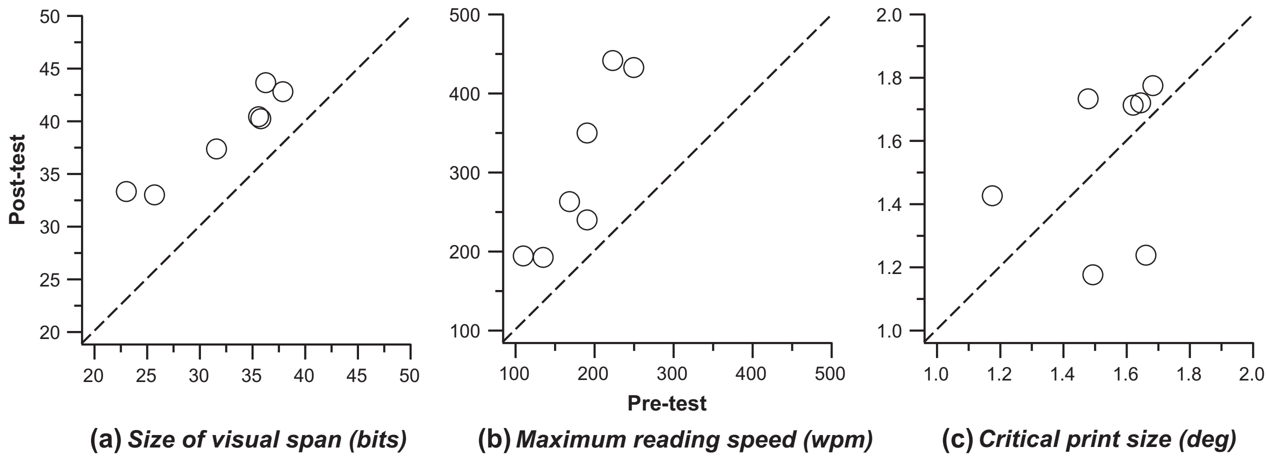

Fig. 5. Comparisons of (a) the size of the visual span (bits), (b) the maximum reading speed (wpm) and (c) the critical print size (deg) between pre- and post-tests. The dashed line in each panel represents the 1:1 line, indicating no change in value between pre- and post-test. For the size of the visual span and maximum reading speed, the post-test values are larger than those at pre-tests, indicating an improvement following training.

Table 2
Changes in the size of the visual span and maximum reading speed in the present study and in Chung, Legge, and Cheung (2004).

	Present study (reading trigrams)	Chung, Legge, and Cheung (2004) (random trigrams)
Size of visual span in bits (post–pre difference) (mean ± 95% CI)	6.44 ± 1.53	7.37 ± 0.82
Maximum reading speed (post/pre ratio) (mean ± 95% CI)	1.64 ± 0.19	1.46 ± 0.14

reading trigrams were not statistically different from Chung, Legge, and Cheung (2004) who used random trigrams for training (Table 2). Specifically, the size of the visual span improved by 6.44 ± 1.53 bits and the maximum reading speed improved by 63.6 ± 18.6%. Based on the similar magnitudes of improvement between our study and that of Chung, Legge, and Cheung (2004), we refuted our hypothesis that training using stimuli that observe the frequently occurring regularities in English usage would be more effective in improving reading speed than using random trigrams.

Was our original hypothesis wrong? The primary measures of our study were the visual-span profile and reading speed. In relation to the visual-span profile, Chung, Legge, and Cheung (2004) used the same procedures and stimuli (random trigrams) to measure the visual-span profile before, during and after training. Here, we used random trigrams for the pre- and post-test measurements, but not for training. Yet, we found a similar magnitude of improvement as Chung, Legge, and Cheung (2004). This result suggests that the visual span is limited by the bottom-up process of letter recognition, as originally proposed by Legge, Mansfield, and Chung (2001), because the improvement following training does not depend on whether the trigrams used during training are uncorrelated random letters, or correlated bigrams and/or trigrams that are found in everyday English usage. It should be noted that not only was the change in the size of the visual span between pre- and post-test similar for the current study and Chung, Legge, and Cheung (2004), the change in the size of the visual span across training sessions was also similar between the two studies, again implying that the visual-span profile is limited by the bottom-up process of letter recognition.

Fig. 3 shows that there is a significant improvement in the size of the visual span between the pre-test and the first training session, averaging 4.51 ± 1.37 bits. However, such a significant improvement was also found in Chung, Legge, and Cheung (2004). A reanalysis of their data only for the group of observers trained in the lower field revealed that the size of the visual span increased from 31.27 ± 3.91 bits at pre-test to 34.45 ± 4.00 bits at the first training session, representing an average increase of 3.18 ± 1.38 bits (paired *t*-test: $t_{(df=5)} = 4.51$, $p = 0.006$). Given that a large increase in the size of the visual span occurred on the first training session regardless of the type of trigrams used (reading trigrams vs. random trigrams), the significant increase cannot be attributed to the different types of trigram stimuli used in the pre-test and the first training session in our study. Rather, the improvement might represent some general learning of the observers, for example, when to pay attention, or how to perform the task in general. A common way to factor out the general improvement due to merely performing the same task twice is to compare the improvement of trained observers with that of observers who did not receive any training. We did not include a no-training control group in this study; however, Chung, Legge, and Cheung (2004) included such a group. In their study, observers in the no-training group received only pre- and post-tests that were separated by four days, with no intervening training. Across observers, the no-training control group performed better on both the visual span and reading speed measurements on post-test, compared with pre-test, but the differences were not statistically different. This implies that the significant improvement in the size of the visual span following the first training session is likely to be due to genuine perceptual learning, instead of observers simply learning when to pay attention or how to perform the task in general.

An alternative explanation for the similar magnitudes of improvement found in the present study and in Chung, Legge, and Cheung (2004) is that our method of extracting reading trigrams from our set of sentences used for the reading task might have included many trigrams that were neither real words nor pseudo-words (strings of letters that are not words but are orthographically and phonologically word-plausible), such as "*ssa*", "*tte*", "*gsi*". There is evidence that letter identification within letter strings is more accurate and faster for words and pseudo-words than for non-word letter strings – the word superiority and pseudo-word superiority effects (Coch & Mitra, 2010; Grainger & Jacobs, 2005; Proverbio, Vecchi, & Zani, 2004; Reicher, 1969). These advantages of words and pseudo-words over non-words are usually attributed to a high-level top-down mechanism (McClelland & Rumelhart, 1981). Indeed, these word and pseudo-word superiority effects were the basis of our hypothesis that reading trigrams might be a more effective training stimulus than random trigrams. In Chung, Legge, and Cheung (2004), because the trigrams used for training were random sequences of letters, the majority of them would be expected to be non-word letter strings. In our study, if the majority of the reading trigrams used for training were also non-word random letter strings instead of words or pseudo-words, then it would be of no surprise that the effectiveness of our training paradigm was so similar to that of Chung, Legge, and Cheung (2004). An analysis of the reading trigrams used during training in our study showed that 79% of the reading trigrams were words or pseudo-words,[5] while a similar analysis showed that only 17% of the trigrams used during training in Chung, Legge, and Cheung (2004) were words or pseudo-words. Despite the much larger proportion of word and pseudo-word trigrams used during training in our study, compared with Chung, Legge, and Cheung (2004), the effectiveness of the training paradigms appeared to be the same, providing further support for the bottom-up processing of letter recognition as a limitation on the visual span.

In relation to reading speed, Legge et al. (2001, 2007) postulated that visual span is the bottleneck on reading speed. If so, then based on our result that the change in the size of the visual-span profile was similar regardless of whether reading trigrams or random trigrams were used for training, we would expect that the change in reading speed that accompanied the change in visual span would also be similar between the current study and Chung, Legge, and Cheung (2004). In contrary, if our training protocol improved observers' ability to recognize frequently occurring bigrams or trigrams (letters that are correlated with one another), or the ability to holistically process chunks of letters in real words, then the improvement in reading speed following training using reading trigrams would be expected to be larger than training using random trigrams. Our result shows that the former case was true, consistent with the notion that visual span, which is limited by the bottom-up process of letter recognition, is the bottleneck on reading speed.

A piece of evidence illustrating that our training method did not specifically improve observers' ability to recognize reading elements (frequently occurring letters, bigrams or trigrams in English usage) is shown in Fig. 6. This figure shows the difference in performance accuracy between pre- and post-test for each letter '*a*'–'*z*', as a function of the average number of presentation of the letter during training and pre- and post-tests. Results are plotted separately for the left, middle or right position of a trigram. Clearly, the improvement in performance accuracy for a given letter does not depend on the number of presentation of the letter, nor does it depend on the pre-test performance for identifying the letter. For instance, as the middle letter of a trigram, the letter '*e*' was presented 729 ± 20 times and the letter '*j*' only 44 ± 5 times during training. Their pre-test recognition values were very similar (48 ± 11% and 49 ± 23%), but the post-test recognition value was only 59 ± 18% for '*e*' and 74 ± 15% for '*j*'. In other words, the efficacy of our training on letter recognition does not depend on the number of times a letter was presented during training. Our results, therefore, imply that any improvement in letter recognition following our training method is likely to be the result of a general letter recognition improvement and is not specific to the particular letter, or combinations of letters presented during training. Previously, Huckauf and Nazir (2007) showed that observers improved only on the set of trigrams presented during training, with very little transfer of improvement to the recognition of letters in trigrams that were not presented during training. However, in their study, the set of trigram stimuli used was very small. Therefore, what their observers learned during training could be the specific combinations of letters, or even the specific stimuli, instead of a general and genuine letter recognition improvement. As such, there was very little

[5] We used a pseudoword generator (Keuleers & Brysbaert, 2010) to classify non-word trigrams as pseudo-words or not. It allows for the generation of written pseudowords based on the orthographic definitions of syllables for a given language.

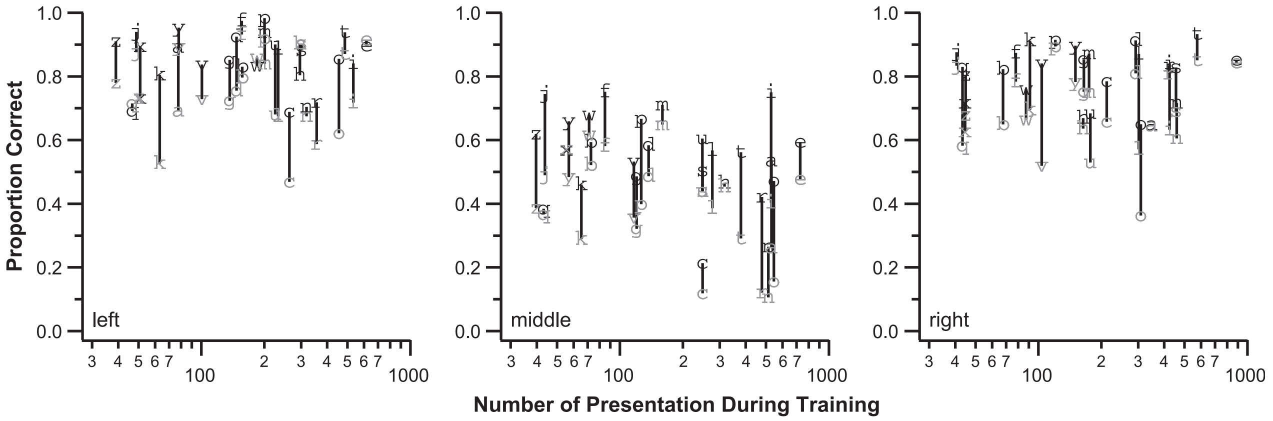

Fig. 6. The improvement in letter recognition performance between the pre- and post-test is plotted for each of the 26 letters of the alphabet, as a function of the number of presentation of each letter. Data are plotted separately when the letters occupied the left, middle or the right letter slot of a trigram. The vertical line connecting a given letter in each panel represents the difference in performance accuracy between the pre- (plotted in gray) and post-tests (plotted in black).

transfer of improvement to untrained trigram stimuli in their study.

What is the "general improvement in letter recognition"? To recognize letters, we need to first detect the letter features and then integrate them to form a percept (Pelli et al., 2006). It has been suggested that the feature integration process is erroneous in the periphery when letters are in close proximity to each other, as in this study. This phenomenon, called crowding, has been attributed to an excessive integration of features from the target and its flanking elements (Levi, 2008; Pelli, Palomares, & Majaj, 2004), and thus could severely limit our ability to recognize letters. Crowding has also been suggested as a major limiting factor on the size of the visual span (Legge et al., 2007) and reading speed in the periphery (Pelli et al., 2007). Recently, Sun, Chung, and Tjan (2010) showed that the mechanism underlying the reduction of crowding following training on identifying crowded letters is attributable to the perceptual window being more capable of adjusting its size to gather relevant input from the object of interest and its flankers. In other words, the feature integration process becomes more efficient with training. We expect that this is indeed what underlies the general improvement in letter recognition for our observers.

Several caveats should be kept in mind when evaluating our findings. First, in normal English usage, many reading trigrams have specific relative locations within words. For instance, trigrams "*ing*" and "*ion*" almost always occur at the end of a word and these words are usually relatively long, thus pushing the usual location of these trigrams toward the right visual field. In our study, each reading trigram presented during training was randomly presented at any of the letter position within six letter slots left and right of the vertical midline. Therefore, some of the trials might not have helped observers in recognizing common reading trigrams that are more often found in other letter positions. A more effective way to train observers to learn the regularities in combination of letters in English usage might be to present reading trigrams according to their usual locations in English words. Future studies may need to consider the locations of (combinations of) letters as they often appear in English words. Second, recognizing letters within a trigram is intrinsically different from recognizing letters within most English words because trigrams comprise only three letters. If crowding is an important limiting factor on letter recognition and that the underlying mechanism for improvement in observers' performance for identifying trigrams is an improved efficiency in the feature integration process (see above), then our trigrams might not have been the most effective stimuli as only the middle letter of each trigram is crowded by letters on both the right and left. A more effective stimulus could be pentagrams (sequences of five letters), however, correctly reporting the five letters of each pentagram is very taxing on observers (Ortiz, 2002).

To conclude, we confirm that perceptual learning using a trigram letter recognition task in the periphery leads to an accompanied improvement in reading speed. The effectiveness of the trigram letter recognition task is similar whether the trigrams are reading trigrams that respect the regularities of how letters are combined in common English usage, or random trigrams, implying that what observers learned during training was some general aspects of letter recognition.

Acknowledgments

This research was supported by research Grant R01-EY012810 from NIH. We thank Daniel Coates and Girish Kumar for their helpful comments on an earlier version of this manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.visres.2012.06.012.

References

Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: Evidence from ERPs. *Neuroreport, 15*, 545–548.

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

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

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 443–446.

Cheong, A. M. Y., Legge, G. E., Lawrence, M., Cheung, S.-H., & Ruff, M. (2007). Relation between slow visual processing and reading speed in people with macular degeneration. *Vision Research, 47*, 2943–2955.

Cheung, S. H., Kaliie, C. S., Legge, G. E., & Cheong, A. M. (2008). Nonlinear mixed-effects modeling of MNREAD data. *Investigative Ophthalmology and Visual Science, 49*, 828–835.

Chung, S. T. L., Legge, G. E., & Cheung, S. H. (2004). Letter recognition and reading speed in peripheral vision benefit from perceptual learning. *Vision Research, 44*, 695–709.

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.

Coch, D., & Mitra, P. (2010). Word and pseudoword superiority effects reflected in the ERP waveform. *Brain Research, 1329*, 159–174.

Elliott, D. B., Trukolo-Ilic, M., Strong, J. G., Pace, R., Plotkin, A., & Bevers, P. (1997). Demographic characteristics of the vision-disabled elderly. *Investigative Ophthalmology and Visual Science, 38*, 2566–2575.

Fahle, M., & Henke-Fahle, S. (1996). Interobserver variance in perceptual performance and learning. *Investigative Ophthalmology and Visual Science, 37*, 869–877.

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59*, 167–192.

Grainger, J., & Jacobs, A. M. (2005). Pseudoword context effects on letter perception: The role of word misperception. *European Journal of Cognitive Psychology, 17*, 289–318.

Huckauf, A., & Nazir, T. (2007). How odgcrnwi becomes crowding: Stimulus-specific learning reduces crowding. *Journal of Vision, 7*, 1–12.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*(3), 627–633.

Legge, G. E. (2007). *Psychophysics of reading in normal and low vision*. Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Legge, G. E., Cheung, S.-H., Yu, D., Chung, S. T. L., Lee, H.-W., & Owens, D. (2007). The case for the visual span as a sensory bottleneck in reading. *Journal of Vision, 7*, 1–15.

Legge, G. E., Mansfield, J. S., & Chung, S. T. L. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research, 41*, 725–734.

Levi, D. M. (2008). Crowding – An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*, 635–654.

Mandelbaum, J., & Sloan, L. L. (1947). Peripheral visual acuity. *American Journal of Ophthalmology, 30*, 581–588.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375–407.

Ortiz, A. (2002). Perceptual properties of letter recognition in central and peripheral vision. Unpublished doctoral thesis, University of Minnesota.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research, 46*, 4646–4674.

Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature, 423*, 752–756.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*, 1136–1169.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision, 7*, 1–36.

Proverbio, A. M., Vecchi, L., & Zani, A. (2004). From orthography to phonetics: ERP measures of grapheme-to-phoneme conversion mechanisms in reading. *Journal of Cognitive Neuroscience, 16*, 301–317.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*, 275–280.

Rubin, G. S. (2001). Vision rehabilitation for patients with age-related macular degeneration. *Eye, 15*(3), 430–435.

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

Rubin, G. S., & Turano, K. (1994). Low vision reading with sequential word presentation. *Vision Research, 34*, 1723–1733.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*, 379–423. 623–656.

Sun, G. J., Chung, S. T. L., & Tjan, B. S. (2010). Ideal observer analysis of crowding and the reduction of crowding through learning. *Journal of Vision, 10*, 1–14.

Wertheim, T. (1980). Peripheral visual acuity. *American Journal of Optometry and Physiological Optics, 57*, 915–924.

Westbury, C., & Buchanan, L. (2002). The probability of the least likely non-length controlled bigram affects lexical decision reaction times. *Brain and Language, 81*, 66–78.

Westheimer, G. (1979). The spatial sense of the eye. Proctor lecture. *Investigative Ophthalmology and Visual Science, 18*, 893–912.

Weymouth, F. W. (1958). Visual sensory units and the minimal angle of resolution. *American Journal of Ophthalmology, 46*, 102–113.

Yu, D., Legge, G. E., Park, H., Gage, E., & Chung, S. T. L. (2010). Development of a training protocol to improve reading performance in peripheral vision. *Vision Research, 50*, 36–45.