Learning to recognize letters in the periphery: Effects of repeated exposure, letter frequency, and letter complexity

Jesse S. Husk   College of Optometry, Ohio State University, Columbus, OH, USA ✉

Deyue Yu   College of Optometry, Ohio State University, Columbus, OH, USA 🏠 ✉

Patients with central vision loss must rely on their peripheral vision for reading. Unfortunately, limitations of peripheral vision, such as crowding, pose significant challenges to letter recognition. As a result, there is a need for developing effective training methods for improving crowded letter recognition in the periphery. Several studies have shown that extensive practice with letter stimuli is beneficial to peripheral letter recognition. Here, we explore stimulus-related factors that might influence the effectiveness of peripheral letter recognition training. Specifically, we examined letter exposure (number of letter occurrences), frequency of letter use in English print, and letter complexity and evaluated their contributions to the amount of improvement observed in crowded letter recognition following training. We analyzed data collected across a range of training protocols. Using linear regression, we identified the best-fitting model and observed that all three stimulus-related factors contributed to improvement in peripheral letter recognition with letter exposure being the most important factor. As an important explanatory variable, pretest accuracy was included in the model as well to avoid estimate biases and was shown to have influence on the relationship between training improvement and letter exposure. When developing training protocols for peripheral letter recognition, it may be beneficial to not only consider the overall length of training, but also to tailor the number of stimulus occurrences for each letter according to its initial performance level, frequency, and complexity.

Introduction

Reading, a daily visual task, is heavily reliant upon foveal vision. In the periphery, reading performance declines notably with eccentricity (Legge et al., 2001). It has been shown that reading and letter recognition are limited by three sensory factors (Yu et al., 2014): resolution (acuity), crowding, and mislocation. Acuity, the smallest letter size required to accurately identify isolated letters, relates generally to the observer's spatial resolution for fine details and decreases with eccentricity in a linear relationship (Anstis, 1974). When multiple letters are presented simultaneously, as often occurs in print, the nearby presence of neighboring letters can reduce the ability to recognize the target letter even when letter size is well above the acuity limit. This phenomenon is known as crowding, the effect of which is minimal at the fovea and increases with eccentricity (Pelli et al., 2004; Whitney & Levi, 2011). Mislocations are position errors due to uncertainty about the relative position of correctly recognized letters. These errors also become greater with eccentricity (Yu et al., 2014). He, Legge, and Yu (2013) investigated the degree to which peripheral reading is limited by these three factors and found that, although all factors play a role in determining performance, crowding appears to be the primary limiting factor in peripheral reading.

People with central vision loss often complain of reading difficulty using their residual vision (Elliott et al., 1997; Mitchell & Bradley, 2006). Improving reading performance in the periphery is, therefore, important for the rehabilitation of these patients. Perceptual learning–based training interventions have been found to successfully enhance reading performance in the normal periphery (e.g., Yu, Cheung et al., 2010; Yu, Legge et al., 2010) and in patients with central vision loss (e.g., Chung, 2011; Tarita-Nistor et al., 2014). Perceptual learning generally involves extensive practice of a perceptual task and results in long-term modification of perception and behavior (Fahle & Poggio, 2002; Gibson, 1963). One of the advantages of employing perceptual learning is that the training method itself can be systematically studied and

optimized for performance improvements. To develop an effective training protocol, we need to consider various components of the protocol, including the stimuli, task, duration, and frequency of training. The goal of this study is to evaluate the roles of stimulus-related factors in perceptual learning regimens for peripheral reading. We made use of perceptual learning data collected across two experimental contexts (Yu, 2013; Yu, Legge et al., 2010) and selected four training groups and two no-training control groups for our analyses. Although the details of the training procedure differed across the four training groups, the experiments shared two main common design components that enabled us to directly compare learning across groups: (a) In all cases, training stimuli were trigrams (strings of three letters) and (b) pretest and posttest performance was measured on a trigram letter recognition task for all groups. This shared methodological structure allowed us to investigate common stimulus-related factors that influence learning irrespective of the specific training task context.

Despite letters being well-learned, familiar patterns, peripheral viewing can pose significant challenges to accurate letter recognition. We hypothesized that under the more taxing condition, differential learning of individual letters might emerge on the basis of factors that varied across the different letters. Specifically, we examined three stimulus-related factors that might impact the effectiveness of training for peripheral letter recognition under crowded conditions: letter exposure, frequency of letter use in English print, and spatial complexity. Letter exposure refers to the number of occurrences of the letter during training and often differs in different training studies as determined by the study design. In some training conditions, certain letters may be presented more often than others. All these variations enable us to examine the impact of repeated letter exposure on learning. Letter frequency, frequency of letter use in English print, has shown little impact on letter identification accuracy in native English speakers (Appelman & Mayzner 1981; Mason 1982). To our knowledge, the letter frequency effect has not been previously studied in the context of learning to improve crowded letter recognition in peripheral vision. Although we do process text information presented outside of the fovea (parafoveal processing) and make use of the extracted information (Schotter et al., 2012), identifying crowded letters specifically using peripheral vision is still a demanding task. Many learning studies have shown that training utilizing short-term exposure to letter stimuli can help enhance peripheral letter recognition (e.g., an average increase of 0.08 to 0.11 in accuracy in Yu, Legge et al., 2010, and 0.16 to 0.17 in Yu, 2013) despite already acquired years of letter reading experience through central vision. It is possible that the already existing long-term letter exposure has

an impact on the amount of improvement for peripherally presented letters. Spatial complexity of letters has previously been reported as an important factor in crowded letter recognition (Bernard & Chung, 2011; Yu, 2015). Higher target complexity corresponds to fewer errors in identifying the target letter compared to lower target complexity (Yu, 2015). Higher complexity, however, may not be an advantage in perceptual learning because more complex letters likely have more features to be learned, which may require more practice than learning less complex letters.

In the present paper, we evaluated the impact of three stimulus-related factors on learning of crowded letter recognition in the periphery by modeling letter-by-letter improvements in performance following training as a function of letter exposure, letter frequency in the English print, and spatial complexity of the letters.

Methods

Experimental procedures

We analyzed the data collected in two previous studies that focused on using perceptual learning to enhance reading speed in the periphery (Yu, 2013; Yu, Legge et al., 2010). Yu, Legge et al. (2010) investigated several character-based training tasks, including a trigram letter recognition task and a lexical decision task, and compared the extent to which each task could improve reading speed in peripheral vision with extensive training. Yu (2013) developed nontask-based training methods utilizing repeated stimulus exposure and priming of stimulus identity. Although the two studies focused on different research questions and have some differences in training details, the experimental procedures adopted to present the stimuli and to operationalize learning and testing were fundamentally identical. The present study examined only the training groups that were trained on stimuli composed of three letters. Four training groups, the trigram letter recognition group and the lexical decision training group from Yu, Legge et al. (2010) and both nontask-based training groups from Yu (2013) were included in our analyses. Two no-training control groups, one from each study, were also included. There were a total of six groups. Each group comprised seven young adults with normal or corrected-to-normal vision.

All testing was conducted binocularly in a dimly lit room at a viewing distance of 40 cm. Subjects were asked to maintain stable fixation at a dot in the center of the display. Across all six groups, the testing and training stimuli were always trigrams presented at different letter distances to the left and right of the

midline at 10° in the lower visual field. All characters were lowercase English letters and rendered as black characters on a white background. A monospaced Courier font and the standard letter spacing (1.16 × x-width) were used. As shown in Figure 1, the position of the trigram (denoted by the position of the middle letter of the trigram) ranged from −5 to 5 (a total of 11 positions). Positive values refer to the positions to the right of the midline; negative values signify the opposite. Position 0 corresponds to the location directly below fixation at the midline. The stimulus presentation duration was 106 ms in Yu (2013) and 105 ms in Yu, Legge et al. (2010).

Print size (defined as x-height) was 2.5° in Yu (2013) and 3.5° in Yu, Legge et al. (2010)—both larger than the critical print size (CPS) at the testing location. CPS refers to the smallest print size above which reading speed reaches plateau. At 10° eccentricity in the lower visual field, CPS has always been found to be smaller than 2.5° for normally sighted young adults (e.g., Chung et al., 1998). Because trigram letter recognition shows the same qualitative dependence on print size as reading speed (Legge et al., 2007), print sizes of 2.5° and 3.5° (both above the CPS) provide similar levels of performance in trigram letter recognition. In short, our analyses and findings should not be affected by the use of different print sizes in the two studies.

Pretest and posttest

All subjects completed a pretest and a posttest that were 1 week apart. The posttest was always conducted on the next day after the last training session for the training groups. Although the task batteries that served as the pretests and posttests in the two studies comprised different combinations of tasks, both studies included a trigram letter recognition task at 10° below fixation. In the trigram task, the stimuli were random letter trigrams (random strings of three letters). Subjects were instructed to report all three letters from left to right. There were 20 trials per position and a total of 220 trials.

Besides the usage of different print sizes mentioned earlier, the only other procedural difference between the two studies was the letter sampling method in the pretest and posttest measurements of trigram letter recognition. To construct a trigram, letters were selected with replacement in Yu, Legge et al. (2010) but without replacement (no letter repetition within trigrams) in Yu (2013). Only 12% of the trials in Yu, Legge et al. (2010) had letters repeated within the trigram. We have excluded these trials from our data analysis in the present study.
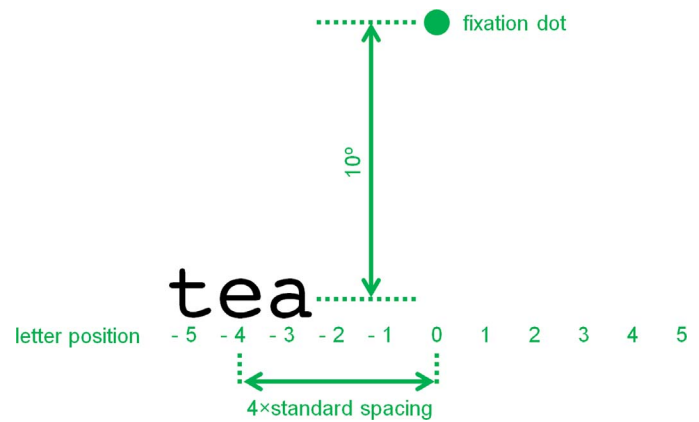

Figure 1. A sample trigram stimulus. In this example, the trigram "tea" is located at position −4. The three letters are positioned at −5 ("t"), −4 ("e"), and −3 ("a"), respectively. Position 0 is right below fixation. Position −4 indicates that the corresponding letter is to the left of the vertical midline and that the distance between the letter and the midline, or position 0, is four times the standard letter spacing.

Training

Two of the six groups analyzed in this study, the trigram letter recognition group and the lexical decision group from Yu, Legge et al. (2010), received explicit task-based training between the pretest and posttest. For both groups, training occurred in daily 1-hr sessions on four consecutive days. In the trigram letter recognition group, subjects practiced on the trigram letter recognition task for a total of 3,520 trials (880 trials per day). The lexical decision training group completed 5,400 trials of a lexical decision task with 1,350 trials per day. In each trial, subjects were instructed to report whether the trigram was a word or nonword.[1] Trigrams were words for half of the trials.

Two nontask-based training groups from Yu (2013), the trigram with repetition group and the trigram without repetition group, were included in the present study. For both groups, training consisted of viewing trigrams on five successive days. Each training session had 26 blocks, one for each letter. Within each block, the middle letter of the trigram (target) was always the same on each trial. Target identity was revealed to the subjects in advance. The left and right letters were randomly selected with replacement across trials. Subject had no task but rather was instructed to attend to each trial and learn to recognize the known target letter. The with-repetition group had the option of repeating a trigram stimulus for as many times as preferred. The without-repetition group did not have this option. On average, the with-repetition group viewed 2,812 more trigrams than the without-repetition

group who viewed a total of 7,150 trigrams (275 trigrams per letter[2]).

The common component across the four training designs was that the stimuli were always trigrams. The key differences were in the stimulus exposure (total number of trigrams) and in the process of selecting letters to compose a trigram (reflected as exposure frequencies of individual letters). Depending on the method of letter sampling, the number of occurrences across individual letters may be the same or similar or may differ greatly across individual letters.

Statistical analyses

Linear regression modeling was used to examine the effects of letter exposure, the frequency of letter use in English print, and letter complexity on performance improvement for crowded letter recognition. We also identified pretest accuracy as another important explanatory variable and included it in the model to avoid producing biased estimates.

Outcome variables: Performance improvement

The outcome variable was performance improvement, calculated as the difference in letter recognition accuracy between the posttest and pretest. We considered only the letters presented at the middle position in the trigram as the middle letter is the most crowded among the three letters of the trigram (Yu et al., 2014). These analyses were also repeated for recognition of all three letters of the trigram, and the results were qualitatively the same (see supplemental results). In our analyses, the unit of observation was the individual alphabet letter observed in an experimental group. In other words, different observations differed in regard to letter identity (26 alphabet letters) and experimental group (six groups). As a result, there were a total of 156 observations with each observation based on trials cumulated across 11 letter positions and seven subjects within the group. Data accumulation across letter positions and subjects was necessary for obtaining accurate estimates of letter recognition accuracy and performance change. Because there is a dissociation between crowding and mislocation (Yu et al., 2014; Xiong et al., 2015) and our interest is in crowding, we excluded the trials with target misplacement errors (one type of mislocation errors). In the removed trials (about 5% of the total trials), the middle letter of the trigram was identified correctly but reported as being in one of the two flanking positions. As shown in Supplementary Tables S4 and S5, the exclusion of these trials had no significant impact on the resulting model.

	Mean	Standard deviation	Minimum	Maximum
Improvement (%)	12.99	14.46	−16.79	60.80
Pretest accuracy (%)	54.32	18.94	14.55	95.24
Letter exposure	167	199	0	1230
Letter frequency (%)	3.85	3.31	0.09	12.55
Letter complexity	152	24	118	194

Table 1. Summary statistics for key variables in the regression modeling.

Stimulus-related factor: Letter exposure

For each letter, letter exposure refers to repeated letter occurrences during the training sessions. An aggregate measure of the number of occurrences, averaged across subjects within each group, was obtained for each letter by taking into account only the middle letters of the trigrams and counting cumulatively across training days. The trials in the pretest and posttest were not included in the exposure count. The two control groups did not receive any training and therefore were treated as having zero letter exposure between the pretest and posttest. Number of letter occurrences during training ranged from zero to 1,230 (see summary statistics in Table 1).

Stimulus-related factor: Letter frequency

Although it is impossible to determine the lifetime letter exposure for the subjects, we can estimate the relative lifetime exposures to different letters from typical usage of English letters in print. We adopted the lowercase letter frequency measured by Jones and Mewhort (2004) who based their estimation on the frequency of letter use in *The New York Times*. Letter frequency was calculated as the percentage of uses of each letter relative to the total counts for the entire alphabet. For example, the letter "e" occurred 7,741,842 times out of a total of 61,676,894, which yielded a percentage use of 12.55%. Across the 26 letters, letter frequencies ranged from 0.09% to 12.55% with a mean of 3.85% (see Supplementary Table S1 for the list of letter frequencies).

Stimulus-related factor: Letter complexity

Letter complexity, the spatial complexity of the letter stimulus, has been quantified in a variety of ways, including perimetric complexity (Attneave & Arnoult, 1956; Pelli et al., 2006), length of morphological skeleton (i.e., total stroke length; Bernard & Chung, 2011), and stroke frequency (Majaj et al., 2002; Zhang et al., 2007). Complexities computed with different methods are highly correlated (Wang et al., 2014). In the present study, we employed the length of the

morphological skeleton to quantify the letter complexity because this complexity measure has been previously used to investigate letter recognition performance under crowded conditions (Bernard & Chung, 2011). The complexities were calculated using custom Matlab software for letters in Courier font with an x-height of 43 pixels and ranged from 118 to 194 with a mean of 152.

Pretest accuracy

Pretest accuracy varies across letters. A higher level of initial performance constrains the potential for improvement. For instance, the initial recognition accuracy was about 82% for letter "f" and 47% for letter "b." This means that letter "b" has more room to improve and is likely to show greater improvement than letter "f" because of initial performance alone. A preliminary data analysis confirms that there is a significant negative correlation between the pretest accuracy and the post–pre improvement in letter recognition ($r = -0.63$, $p < 0.0001$). The observation of the negative correlation is in agreement with many other perceptual learning studies (Astle et al., 2013; Baldassarre et al., 2012; Fahle & Henke-Fahle, 1996; Tarita-Nistor et al., 2014). Statistically, it is important to include such a significant explanatory variable in the model to avoid constructing an underspecified model and producing biased estimates. Additionally, it was necessary to take into account the differences in the pretest accuracy across letters because improvements can be quantitatively the same but qualitatively different depending on the initial levels of performance (e.g., a change from 50% to 55% vs. a change from 90% to 95%).

Modeling approach

The overall objective was to model performance change as a function of letter exposure, frequency, and complexity after controlling for pretest performance. Prior knowledge of the response curve was used to serve as a guide in choosing the most appropriate functional form for each variable. Logarithmic transformation was performed on two independent variables that have nonlinear relationships with the dependent variable. First, the time course of performance improvement in perceptual learning tends to follow a nonlinear function where fast gains emerge at the beginning of training with diminishing returns over time (Dosher & Lu, 2007). Using the training data from the trigram letter recognition group (Yu, Legge et al., 2010), we assessed whether the same relationship is true for letters. We examined three types of functions—linear, exponential, and logarithmic—and found that the logarithmic function corresponds to the lowest mean squared error and is the best candidate to model the learning curve for crowded letter recognition. Previous data (Yu, 2013; Yu et al., 2010) also indicated a possible nonlinear relationship between the total improvement and pretest accuracy. Comparing among the three types of functions, fitting with a logarithmic function consistently led to the lowest mean squared errors. In the present study, we adopted the logarithms of letter exposure and pretest accuracy to evaluate the nonlinear relationships while preserving the linear model. Base 10 was used for convenience so that the value of two corresponds to 100 occurrences for letter exposure or 100% accuracy for pretest performance. Because the logarithm of 0 is undefined, we recorded instances in which letter exposure was zero to one before performing the logarithmic transformation.

We were interested in examining both the linear and interaction terms in the model. The selection of the most parsimonious model was accomplished via a backward elimination strategy. The starting model included the four predictors of interest and all possible interaction terms. Nonsignificant terms were progressively eliminated, one at a time, starting with the four-way interaction, followed by the three-way and two-way interactions and the main effect terms. For significant interaction terms, the corresponding main effect terms remained in the model regardless of their statistical significance.

Bootstrapped confidence interval

Bootstrapping, a distribution-free method, was used to obtain the confidence intervals for the regression estimates. Bootstrapping confidence intervals is advantageous in cases where the sample size is relatively small such that the conditions of the Central Limit Theorem may be violated, and normal distribution assumptions may not apply (Fox, 2015). Although there is no absolute standard for what constitutes a small sample size, we opted to use bootstrapped confidence intervals to avoid making assumptions about the underlying distribution of the data. To generate the bootstrapped confidence intervals, 5,000 iterations of random sampling with replacement were used. At each iteration, the resampled observations were fit using regression, and the resulting coefficients were recorded. After 5,000 iterations, the resulting distribution of coefficient values was examined, and the 95% confidence interval was obtained for each coefficient in that stage of the model (Fox, 2015). When the 95% confidence interval for a coefficient included zero, the corresponding predictive term was considered a nonsignificant predictor of the outcome.

Results

Across all groups, recognition accuracy of the middle letter of the trigram improved an average of 12.99% from pretest to posttest (Table 1). Table 2 shows the significant coefficients retained in the model after the backward selection procedure. The four main effect predictors—letter exposure, letter frequency, letter complexity, and pretest accuracy—show significant associations with the outcome measure. There was also a significant interaction between log(Letter exposure) and log(Pretest accuracy). The best-fitting model relating mean improvement $E(\text{improvement} \mid x)$ to the four explanatory variables is given by the equation:

$$\begin{aligned} E(\text{improvement}) = {} & 87.23 \\ &+ 21.09 \times \log(\text{Letter exposure}) \\ &- 1.17 \times \text{Letter frequency} \\ &- 0.10 \times \text{Letter complexity} \\ &- 35.55 \times \log(\text{Pretest accuracy}) \\ &- 10.14 \times \log(\text{Letter exposure}) \\ &\quad \times \log(\text{Pretest accuracy}) \end{aligned}$$

This model indicates that improvement in crowded letter recognition is significantly related to letter exposure, frequency, and complexity after controlling for pretest accuracy. As shown in Figure 2, more frequent letters are expected to improve less than less frequent letters while holding letter exposure, letter complexity, and initial accuracy constant. More complex letters improve less on average than less complex letters when exposure, frequency, and initial performance are held fixed.

The interaction term suggests that the effect of letter exposure on improvement is greater at lower levels of pretest performance. In other words, the lower the initial accuracy of letter recognition, the greater improvement the letter gains with repeated letter exposure. On the other hand, the more letter exposure, the greater the effect of pretest accuracy on improvement. Note that the contribution of the interaction term to explaining the total variance in the improvement is small ($sr^2 = 0.02$). Overall, $E(\text{improvement})$ is estimated to increase with increasing the number of letter occurrences while controlling for the other factors in the model; increasing the pretest accuracy in letter recognition leads to reduction in the mean performance improvement, holding the other predictor variables constant.

Together, these results suggest that increasing the number of trials in which a subject views crowded letters is associated with greater improvement in letter recognition although the rate of improvement depends on the level of pretest performance. Meanwhile,

Predictive term	Coefficient	95% CI	sr^2
log(Letter exposure)	21.09	[7.57, 35.15]	0.09
Letter frequency	−1.17	[−1.63, −0.76]	0.06
Letter complexity	−0.10	[−0.17, −0.03]	0.02
log(Pretest accuracy)	−35.55	[−49.85, −22.27]	0.36
log(Letter exposure): log(Pretest accuracy)	−10.14	[−18.26, −2.34]	0.02
Intercept	87.23	[63.29, 113.50]	

Table 2. Results of regression model for improvement ($R^2 = 0.59$; $R^2_{\text{adj}} = 0.57$). *Notes*: Dependent variable is middle letter improvement. Predictors are middle letter exposure, letter frequency, letter complexity, and pretest performance. 95% CI: 95% bootstrapped confidence interval. sr^2: the squared semipartial correlation.

improvement is less in the context of letters that are used more frequently in English print or more spatially complex. Higher initial performance level is associated with less improvement with letter exposure modulating the size of the effect of pretest accuracy. Adjusted R square, R^2_{adj}, was chosen as a measure of model adequacy. As shown in Table 2, the model explains 57% of the variability of the improvement. To assess the unique contribution and relative importance of each predictive term in determining $E(\text{improvement})$, we also calculated semipartial correlations. Squared semipartial correlation (sr^2) indicates the amount of decrease in R^2 when the variable is removed from the regression equation (Cohen et al., 2013). Among the three stimulus-related factors, letter exposure contributed the most and letter complexity contributed the least to the explained variation in the improvement.

The analyses presented in Table 2 focused on improvement in recognizing the crowded, middle letter of the trigram. Letter exposure was measured as the count of occurrences of each letter in the middle letter position within the trigram during training. As shown in Supplemental Table S2, modeling performance improvement with three-letter exposure (number of occurrences of each letter across all three letter positions within the trigram) does not substantially alter the results of the regression analysis. The resulting best model is qualitatively identical to the one shown in Table 2. A similar relationship also holds when we evaluate three-letter improvement (i.e., improvement calculated across all three letters of the trigram) and use total letter exposure across all three positions within trigram as a predictor along with letter frequency, complexity, and pretest accuracy (Supplemental Table S3).

The analyses in the present study were performed after excluding trials in which responses involved target misplacement errors. To determine whether the exclusion of these trials had any significant impact on the resulting model (Table 2), we repeated the regression
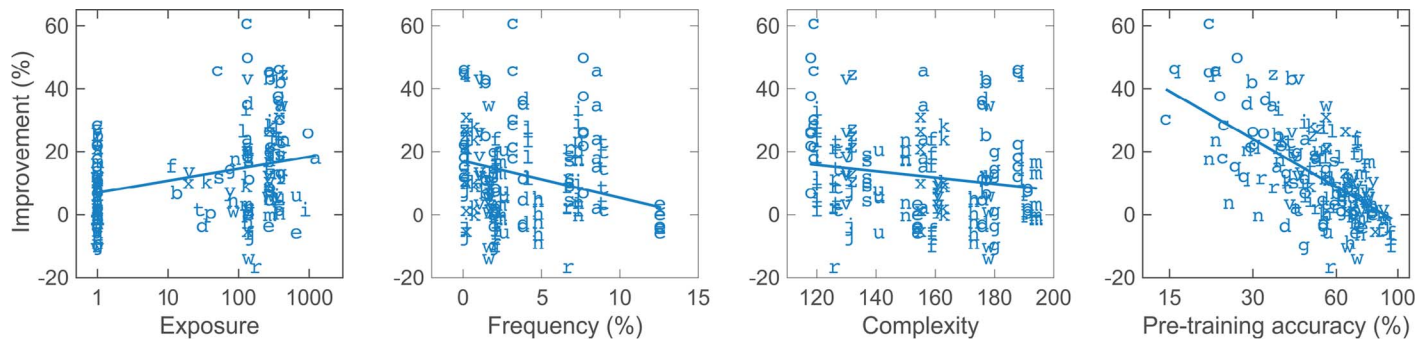
Figure 2. Post–pre improvement as a function of (a) letter exposure, (b) letter frequency, (c) letter complexity, and (d) pretest accuracy. The lines represent the fit to the data given by the model. Each letter represents a data point associated with the letter.

analysis with all trials included. As shown in the Supplemental Tables S4 and S5, target misplacement errors were handled in two different ways: (a) treating the responses as incorrect and (b) treating the responses as correct. We found that inclusion of the trials with target misplacement errors had little influence on the resulting model regardless of error handling approaches.

We tested model stability using a procedure akin to a jackknife procedure in which we repeated the model analysis six times, each time dropping out one of the six experimental groups. The rationale of this analysis was that if an experimental group is causing undue influence on the model, the parameter estimates should vary notably when that group is removed from the analysis. Further, because only one group was dropped at a time, the sample size ($n = 130$) remained large enough to produce stable model solutions with four predictors. Table 3 presents the coefficients obtained during this analysis. The estimate for the interaction term, log(Letter exposure) × log(Pretest accuracy), appears less stable across simulations than the estimates for the main terms. In particular, dropping the trigram letter recognition group or the 2013 control group results in a nonsignificant interaction term. This

is not surprising considering the small contribution of the interaction term to the model ($sr^2 = 0.02$). It suggests that the association between log(Letter exposure) × log(Pretest accuracy) and improvement may be more heavily driven by the trigram letter recognition group and the 2013 control group. As shown in Table 3, regardless of whether the model contains the interaction term, the variability of the coefficient estimates is small relative to the estimates based on all groups.

Discussion and conclusions

We explored three stimulus-related factors (letter exposure, frequency, and complexity) as predictors of the effectiveness of training for peripheral letter recognition while controlling for pretest accuracy. All factors are significantly related to the performance change in crowded letter recognition. Among the three stimulus-related factors, letter exposure contributes the most and the spatial complexity of the letter contributes the least to the explained variation in the performance change. Letters that are initially more

	Group dropped						All groups (with interaction term)	All groups (without interaction term)
Predictive term	2010 control	2010 lexical	2010 trigram	2013 control	2013 without repetition	2013 with repetition		
log(Letter exposure)	26.20	20.79	3.96	3.43	21.53	31.67	21.09	3.81
Letter frequency	−1.26	−0.97	−1.34	−1.20	−1.15	−1.14	−1.17	−1.18
Letter complexity	−0.09	−0.09	−0.07	−0.11	−0.12	−0.11	−0.10	−0.10
log(Pretest accuracy)	−29.05	−36.79	−44.62	−56.81	−34.57	−32.55	−35.55	−52.23
log(Letter exposure): log(Pretest accuracy)	−13.19	−9.68	/	/	−10.58	−16.35	−10.14	/

Table 3. Coefficient estimates in the model stability analysis that examines the relative impact of dropping out a subset of the data. *Notes*: Dependent variable is middle letter improvement. Predictors are middle letter exposure, letter frequency, letter complexity, and pretest performance. All coefficient estimates presented in the table are significantly different from zero. Coefficient estimates based on all groups with and without the interaction term are also presented here as references.
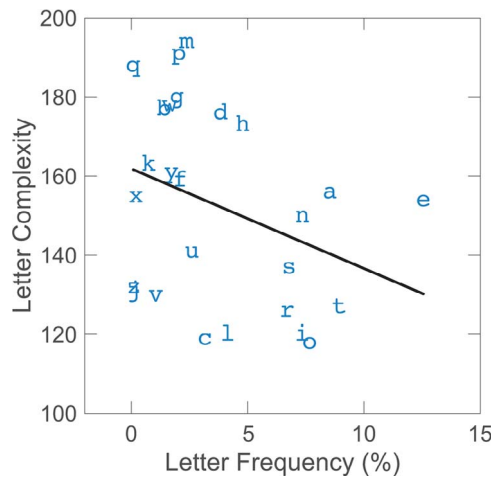
Figure 3. Marginally significant correlation between letter frequency and letter complexity ($r = -0.34$, $p = 0.09$).

difficult to recognize, exposed more frequently during training, used less frequently in English print, or less spatially complex are associated with greater performance gains. Although the effect of letter exposure depends on the level of the pretest performance and vice versa, lesser confidence should be placed in the contribution of this interaction.

Multicollinearity

It is worth noting that frequency and complexity were marginally correlated ($r = -0.34$, $p = 0.09$; Figure 3). High-complexity letters tend to appear less frequently in English print whereas lower-complexity letters span the range of letter frequencies in the language. To evaluate the potential effect of multicollinearity, we examined the variance inflation factors (VIFs; Fox, 2015) to estimate how much the variance of one coefficient is inflated due to linear dependence with other predictors. We found that values of all VIFs ($\leq$ 1.19) were below the common criterion for identifying problematic high VIF. Therefore, we conclude that multicollinearity is not a concern for our models.

Letter exposure

It may not be surprising that greater exposure to training stimuli resulted in greater improvement. However, letters are already familiar objects, and the differences in the amount of exposure to letters during training are very small by comparison to lifetime exposure levels. In the current analysis, the maximum letter exposure is 1,230 trials. We are able to use a logarithmic function to describe the performance improvement as a function of letter exposure (i.e., the

amount of training). What if we extend the training beyond 1,230 trials? The relationship between performance changes and the length of training could vary depending on the time period over which the performance changes are considered. For instance, many studies show that perceptual learning–based improvement plateaus after just a brief period of practice such as a few hours (e.g., Chung et al., 2004; Levi & Polat, 1996). For that reason, most perceptual learning studies only focus their investigations on training over short periods of time. Only a handful of studies have tried to track learning beyond the initial plateau. Li et al. (2008) found that, beyond the initial brief plateau, there could be further substantial improvement with prolonged training, especially for patients with more severe vision loss. It is possible that crowded letter recognition continues to improve with additional letter exposure beyond the range tested here, which raises the question of when the ultimate asymptotic performance can be reached. This has yet to be answered.

Letter frequency

Although it is impossible to estimate lifetime letter exposure levels for our subjects, letter frequency counts in English print can act as a rough proxy for the relative lifetime exposures to different letters. Does long-term exposure matter in the context of letter recognition? Pelli et al. (2006) found that for normal vision, efficiency of letter identification rises quickly with initial learning for both adults learning foreign alphabets and young children learning native alphabets. After the initial period of rapid improvement in efficiency, no further gain was found even with decades of additional reading experience. In other words, human subjects can improve their performance of letter identification to the ceiling level with merely a short-term exposure to letter stimuli. Hence, it is reasonable to observe the lack of letter frequency effect on identification accuracy in native English speakers (Appelman & Mayzner, 1981; Mason, 1982). Although long-term exposure to letter stimuli is typically accrued through central vision for normally sighted people, we found the same result—no correlation between letter frequency and (pretest) recognition accuracy—in the periphery. Interestingly, the present study showed that letter frequency does have an effect on performance improvements with lower-frequency letters associated with more improvement than higher-frequency letters. Does this mean we should allocate more trials to low-frequency letters during the training with a goal of improving reading performance? Practically speaking, when training time is limited, it may not be cost-efficient to do so because low-frequency letters, by definition, appear less frequently in daily reading.

Devoting more training trials to high-frequency letters may not be advantageous either. A previous study found that the improvement through training with trigrams frequently used in the English language did not exceed the improvement following random-letter trigram training (Bernard et al., 2012). Further investigation is needed to evaluate whether or not training paradigms should be optimized based on letter frequency.

Clinical implications

For patients with central vision loss, learning to utilize their residual vision for tasks formerly performed by the fovea is critical for maintaining daily functioning and quality of life. Most people with central vision loss are able to, at least on a certain level, adapt naturally to their vision loss through processes such as recruiting an eccentric retinal location as a replacement for the original fovea (Cheung & Legge, 2005). Training interventions, providing intentional, guided practice under controlled conditions, are important components of low vision rehabilitation complementary to patients' natural adaptations. Developing a training procedure to enhance letter recognition and reading performance in the periphery is especially important for reading rehabilitation in these patients. Although we have had success developing effective training methods for peripheral reading, understanding the roles of various factors in learning can help further improve the efficiency and effectiveness of training. The findings of the present study contribute to our knowledge on how stimulus-related factors and pretest performance might influence performance change for peripheral letter recognition. To optimize training protocols for peripheral letter recognition, it may be beneficial to not only consider the overall length of training, but also to tailor the number of stimulus occurrences for each letter according to its initial performance level, frequency, and complexity. Although we examined only four explanatory factors, we recognize that other factors, such as training task, duration, and frequency of training, may also have significant impacts on the effectiveness of training.

Keywords: letter recognition, peripheral vision, crowding, perceptual learning, low vision

Acknowledgments

Footnotes

[1] To ensure the task was sufficiently difficult to observe learning, a postmask "###" was presented after each trigram, and a shorter stimulus duration (92 ms) was used for two of the subjects. Feedback on performance accuracy was provided to the subjects following each training block.

[2] For the trigram without repetition group, the number of letter exposures was designed to be 275 trials per letter. However, a program error occurred in one training session for one subject, resulting in slightly different numbers of exposures across letters (267, 275, or 283 trials per letter).

References

Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, *14*(7), 589–592.

Appelman, I. B., & Mayzner, M. S. (1981). The letter-frequency effect and the generality of familiarity effects on perception. *Perception & Psychophysics*, *30*(5), 436–446.

Astle, A. T., Li, R. W., Webb, B. S., Levi, D. M., & McGraw, P. V. (2013). A Weber-like law for perceptual learning. *Scientific Reports, 3*, 1–8.

Attneave, F., & Arnoult, M. D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, *53*(6), 452–471.

Baldassarre, A., Lewis, C. M., Committeri, G., Snyder, A. Z., Romani, G. L., & Corbetta, M. (2012). Individual variability in functional connectivity predicts performance of a perceptual task. *Proceedings of the National Academy of Sciences, USA*, *109*(9), 3516–3521.

Bernard, J. B., Arunkumar, A., & Chung, S. T. (2012). Can reading-specific training stimuli improve the effect of perceptual learning on peripheral reading speed? *Vision Research, 66*, 17–25.

Bernard, J.-B., & Chung, S. T. (2011). The dependence of crowding on flanker complexity and target–flanker similarity. *Journal of Vision*, *11*(8):1, 1–16, doi:10.1167/11.8.1. [PubMed] [Article]

Cheung, S.-H., & Legge, G. E. (2005). Functional and cortical adaptations to central vision loss. *Visual Neuroscience*, *22*(2), 187–201.

Chung, S. T. (2011). Improving reading speed for people with central vision loss through perceptual learning. *Investigative Ophthalmology & Visual Science*, *52*(2), 1164–1170. [PubMed] [Article]

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 Research*, *38*(19), 2949–2962.

Chung, S. T., Legge, G. E., Cheung, S. H. (2004). Letter-recognition and reading speed in peripheral vision benefit from perceptual learning. *Vision Research*, *44*(7), 695–709.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Routledge.

Dosher, B. A., & Lu, Z.-L. (2007). The functional form of performance improvements in perceptual learning rates and transfer. *Psychological Science*, *18*(6), 531–539.

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 & Visual Science*, *38*(12), 2566–2575. [PubMed] [Article]

Fahle, M., & Henke-Fahle, S. (1996). Interobserver variance in perceptual performance and learning. *Investigative Ophthalmology & Visual Science*, *37*(5), 869–877. [PubMed] [Article]

Fahle, M., & Poggio, T. (2002). *Perceptual learning*. Cambridge, MA: MIT Press.

Fox, J. (2015). *Applied regression analysis and generalized linear models*. Thousand Oaks, CA: Sage Publications.

Gibson, E. J. (1963). Perceptual learning. *Annual Review of Psychology*, *14*(1), 29–56.

He, Y., Legge, G. E., Yu, D. (2013). Sensory and cognitive influences on the training-related improvement of reading speed in peripheral vision. *Journal of Vision*, *13*(7):14, 1–14, doi:10.1167/13.7.14. [PubMed] [Article]

Jones, M. N., & Mewhort, D. J. (2004). Case-sensitive letter and bigram frequency counts from large-scale English corpora. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 388–396.

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. *Journal of Vision*, *7*(2):9, 1–15, doi:10.1167/7.2.9. [PubMed] [Article]

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 Research*, *41*(6), 725–743.

Levi, D. M., & Polat, U. (1996). Neural plasticity in adults with amblyopia. *Proceedings of the National Academy of Sciences, USA*, *93*(13), 6830–6834.

Li, R. W., Klein, S. A., & Levi, D. M. (2008). Prolonged perceptual learning of positional acuity in adult amblyopia: Perceptual template retuning dynamics. *The Journal of Neuroscience*, *28*(52), 14223–14229.

Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, N. (2002). The role of spatial frequency channels in letter identification. *Vision Research*, *42*(9), 1165–1184.

Mason, M. (1982). More about the letter-frequency effect. *Perception & Psychophysics*, *31*(6), 589–590.

Mitchell, J., & Bradley, C. (2006). Quality of life in age-related macular degeneration: A review of the literature. *Health and Quality of Life Outcomes*, *4*(1), 1–20.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*(28), 4646–4674.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*(12):12, 1136–1169, doi:10.1167/4.12.12. [PubMed] [Article]

Schotter, E. R., Angele, B., & Rayner, K., (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74*(1), 5–35.

Tarita-Nistor, L., Brent, M. H., Steinbach, M. J., Markowitz, S. N., & González, E. (2014). Reading training with threshold stimuli in people with central vision loss: A feasibility study. *Optometry & Vision Science*, *91*(1), 86–96.

Wang, H., He, X., & Legge, G. E., (2014). Effect of pattern complexity on the visual span for Chinese and alphabet characters. *Journal of Vision*, *14*(8):6, 1–17, doi:10.1167/14.8.6. [PubMed] [Article]

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160–168.

Xiong, Y.-Z., Yu, C., & Zhang, J. Y. (2015). Perceptual

learning eases crowding by reducing recognition errors but not position errors. *Journal of Vision*, *15*(11):16, 1–13, doi:10.1167/15.11.6. [PubMed] [Article]

Yu, D. (2013). Improving reading speed in peripheral vision through a non-task-based training. *Investigative Ophthalmology & Visual Science*, *54*(15), 2751–2751. [Article]

Yu, D. (2015). Effect of spatial complexity on resolution, mislocations and crowding. *Journal of Vision*, *15*(12): 99, doi:10.1167/15.12.99. [Abstract]

Yu, D., Cheung, S. H., Legge, G. E., & Chung, S. T. (2010). Reading speed in the peripheral visual field of older adults: Does it benefit from perceptual learning? *Vision Research*, *50*(9), 860–869.

Yu, D., Legge, G. E., Park, H., Gage, E., & Chung, S. T. (2010). Development of a training protocol to improve reading performance in peripheral vision. *Vision Research*, *50*(1), 36–45.

Yu, D., Legge, G. E., Wagoner, G., & Chung, S. T. (2014). Sensory factors limiting horizontal and vertical visual span for letter recognition. *Journal of Vision*, *14*(6):3, 1–17, doi:10.1167/14.6.3. [PubMed] [Article]

Yu, D., Park, H., Gerold, D., & Legge, G. E. (2010). Comparing reading speed for horizontal and vertical English text. *Journal of Vision*, *10*(2):21, 1–17, doi:10.1167/10.2.21. [PubMed] [Article]

Zhang, J. Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2007). Legibility variations of Chinese characters and implications for visual acuity measurement in Chinese reading population. *Investigative Ophthalmology & Visual Science*, *48*(5), 2383–2390. [PubMed] [Article]