Original Article

QJEP

Quarterly Journal of Experimental Psychology
1–15

DOI: 10.1177/17470218221102878
qjep.sagepub.com
SAGE

Lexical competition influences correct and incorrect visual word recognition

Julie Gregg[1], Albrecht W Inhoff[1] and Xingshan Li[2]

Abstract

A growing body of research suggests that visual word recognition is error-prone, and that errors may contribute to inhibitory neighbour frequency effects in word identification and reading. The present study used the neighbourhood frequency effect to examine the relationship between lexical competition and error making during visual word recognition. A novel adaptation of the visual world paradigm (VWP) was used, in which participants selected a briefly presented printed target word from an array containing the target, its higher- or lower-frequency neighbour, an orthographic onset competitor, and an orthographically unrelated distractor word. Analyses of the visual inspection of the arrays suggested that lexical competition occurred when words were correctly identified, as competitors were preferentially viewed as a function of their orthographic similarity with the target, and higher-frequency neighbours were preferentially viewed over lower-frequency neighbours. Orthographic similarity and neighbour frequency also influenced error making. Targets were often mistaken for their neighbours, and these errors were more common for targets with higher-frequency neighbours. The time course of target and neighbour viewing for error trials also provided preliminary evidence for two kinds of errors: early-occurring, perceptual errors and later-occurring selection errors that resulted from unsuccessfully resolved lexical competition. Together, these findings suggest that neighbour frequency effects reflect the contribution of both general lexical competition and occasional errors.

Keywords

Eye movements; visual word recognition; orthographic neighbourhood; lexical competition; visual world paradigm

Received: 15 October 2021; revised: 17 March 2022; accepted: 25 April 2022

Fluent readers recognise words quickly and accurately. In fact, many computational simulations of readers' viewing of text assume that word recognition is accomplished within brief word viewing durations and that word recognition is error free (Reichle et al., 2006; Engbert et al., 2005). Yet, most readers had the experience of confusing one word with another, and those misread words were similar to the actual word (see Rayner et al., 1981).

The intuition that orthographically similar words may be confused has given rise to a large body of research examining the influence of orthographic similarity on visual word recognition. To account for these effects, word recognition models have assumed that the perception of a visual word activates a set of orthographically related lexical candidates, and that this initial phase of lexical uncertainty is subsequently resolved through inhibition that discerns a single candidate (Chen & Mirman, 2012; McClelland & Rumelhart, 1981; the SERIOL model: Whitney, 2001; the spatial coding model: Davis, 2010; the activation verification model: Paap et al., 1982). For these models, competition of candidates is assumed to be the major mechanism of lexical selection, and errors could occur when inhibition yields a candidate other than the target. Therefore, we will refer to these models as activation-inhibition models hereafter.

Inhibitory effects of orthographically similar candidates are well established when a to-be-recognised target word and an orthographically similar word differ by a

[1]Binghamton University, Binghamton, NY, USA
[2]Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing (CAS), China

*Xingshan Li is now affiliated to Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

Corresponding author:
Xingshan Li, Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Road, Beijing 100101, China.
Email: lixs@psych.ac.cn

transposition of two characters (e.g., *clam* and *calm*—referred to as transposition neighbours), and when words differ by the addition or deletion of a character (e.g., *drive* and *dive*), especially when the neighbour has a higher frequency of occurrence than the target (Acha & Perea, 2008; Andrews, 1997; Chambers, 1979; Davis et al., 2009; Johnson, 2009). Length-matched neighbours that differ by a single letter from the target (e.g., *boot* and *boat*—substitution neighbours) show a more complex pattern of effects (see Andrews, 1996, for a review). One example is the influence of neighbour frequency, with higher-frequency neighbours impeding the recognition of the target. While the presence of higher-frequency neighbours is generally inhibitory in transparent orthographies (Carreiras et al., 1997; Grainger, 1990; Grainger & Jacobs, 1996; Grainger et al., 1989; Grainger & Segui, 1990; van Heuven et al., 1998), the effects are inconsistent in opaque orthographies, such as English. One group of studies shows inhibitory effects (Huntsman & Lima, 1996; Paap et al., 2000; Paterson et al., 2009; Perea & Pollatsek, 1998; Sears et al., 2006, Experiment 1A; Slattery, 2009; Warrington et al., 2016; Yao et al., 2021), another shows no effect (Forster & Shen, 1996, Experiment 4; Huntsman & Lima, 2002; Sears et al., 2006, Experiments 1B–3B), and yet another group shows facilitatory effects (Forster & Shen, 1996, Experiments 1–3; Sears et al., 1995; see Slattery, 2009, for a review). Since previous studies did not yield consistent evidence for obligatory inhibition during visual word recognition, it is possible that the occurrence of inhibition depends on lexical and contextual properties. Specifically, it may occur only when the represented form (and meaning) of a competitor has a distinct advantage over the representation of the visible target, for instance, by virtue of being used more often. In this case, the activation of a competitor's form could be strong enough to dominate the representation of the target, thus leading to a word recognition error. Conversely, a relatively weak activation of a neighbour could yield benefits (Chen & Mirman, 2012). Instead of dominating the visible target, it could make the target more distinctive and favour its selection. Eye-tracking studies showed that orthographically similar words affect eye movements during sentence reading. In Pollatsek et al. (1999), target words with many neighbours were skipped (i.e., not fixated) more often than targets with few neighbours, suggesting that increases in the set of lexical candidates facilitated their recognition. The viewing duration of target words that were viewed (fixated) was influenced neither by the size nor by the properties of an orthographic neighbourhood, also suggesting an absence of inhibition. However, targets were occasionally re-read, and this was more common for targets with many than with few neighbours. Moreover, the effect of a large neighbourhood on re-reading was the greatest when the target's highest frequency neighbour was immediately plausible. This was taken as evidence that some targets were initially misread as a neighbour, leading to re-reading when the misread target was anomalous with subsequent context (see also Acha & Perea, 2008). Furthermore, studies have observed inhibitory effects on word reading when the word's substitution or transposed letter neighbour appeared earlier in the sentence; here, priming could have bolstered inhibition and/or the target could have been misread as its transposition neighbour (Pagán et al., 2015; Paterson et al., 2009).

Distributional analyses also suggest that inhibitory effects of orthographic similarity could be the exception rather than the rule. Johnson et al. (2012) fitted the ex-Gaussian distribution to naming latencies, and they observed that the presence of a transposition neighbour influenced the τ but not the μ parameter of the distribution of naming latencies, where τ indexes skewing due to a subset of slow responses and μ indexes the shifts of the distribution means. This pattern of results suggests that neighbours influenced only a subset of trials. In addition, Johnson et al. (2012) observed that misreading occurred more often for words with than without transposition neighbours, and that 86% of these errors reflected partial or complete misarticulation of the word as its neighbour. Thus, inhibitory effects of orthographic similarity could be exception rather than the rule, and they could arise from the misreading of a small subset of words.

The potential for misreading to exert substantial influence on measures of word recognition is also consistent with a growing body of evidence according to which words are occasionally mistaken for their neighbours. Potter et al. (1993) used a Rapid Serial Visual Presentation (RSVP) paradigm in which participants were asked to identify briefly presented words and non-words out-loud. Across several experiments, Potter et al. observed that when pre- or post-target context was presented and that biased towards a substitution neighbour of a target word, participants reported the neighbour in 18% to 34% of trials. In the absence of bias, Paap et al. (2000) observed that lower-frequency words were misread as their higher-frequency neighbours in more than 12% of trials. In studies of silent reading, Slattery (2009) and Warrington et al. (2016) observed inhibition on measures including re-reading for words with higher-frequency substitution neighbours relative to matched controls without, but only when the word's neighbour was a plausible continuation of the sentence, suggesting that words were initially misread as their neighbours (see also Gregg & Inhoff, 2016, Experiments 1 and 2). In an oral reading study, Gregg and Inhoff (2016, Experiment 3) showed that participants misarticulated words as their higher-frequency substitution neighbours in 6% of trials. Together, prior work suggests that lexical competition between orthographic neighbours may result in mistakenly recognising the target as their high-frequency neighbours in a small subset of trials. On these trials, recognition of the target could require an optional and time-consuming verification, or recognition could be erroneous, in particular when the orthographic neighbour has a higher frequency of occurrence than the target.

This study sought to determine the extent to which lexical competition is the exception rather than the rule during visual word recognition. Specifically, it tested the influence of weak and strong orthographic neighbours on the success of visual target selection, and it examined the time course of lexical competition. According to activation-inhibition models of word recognition, strong neighbours should yield larger inhibitory effects than weak neighbours, inhibitory effects should be obtained for both accurate and inaccurate target selections, and inhibition should occur relatively late in the selection process. To gain insight into the generality, nature, and time course of lexical competition, the present study implemented a version of the visual world paradigm (VWP; Cooper, 1974; Tanenhaus et al., 1995) that required the identification of visual rather than spoken words (Meyer & Federmeier, 2008). In a more typical VWP word recognition study, participants listen to a spoken message with a target word, and they select the target from an array of words or pictures, one of which matches the target. The time course of object viewing during the presentation of the array is then used to elucidate the information used for target selection and the time course of information use. Since the task is sensitive to the degree of representational overlap between the target and the objects of the array (see Huettig et al., 2011, for a review) and the temporal unfolding of target selection, it was used in this study to examine the nature and time course of the effects of orthographic neighbours on target selection.

In this study, a to-be-recognised visual target (e.g., *spell*) was briefly presented and then masked. A forced choice array of four words was subsequently presented, and readers were instructed to select the target from it. The array contained four objects: the target word (e.g., *spell*), a higher- or lower-frequency substitution neighbour (e.g., the higher-frequency/strong competitor word *shell*), an onset competitor (e.g., *speed*), and an orthographically unrelated control word (e.g., *trade*). Given the limited processing time for the target word, readers were expected to make occasional target selection errors.

Trials with correct responses were examined to obtain evidence for the generality of inhibitory effects of orthographic neighbours. On strong competition trials, low frequency targets were shown with a higher-frequency neighbour, and on weak competition trials, high-frequency targets were shown with a lower-frequency neighbour. To determine the nature of the time course of the neighbour effect, the time course of neighbour viewing was compared with the time course of onset competitor and control word viewing, a preferential viewing of neighbours over onset competitors and control words suggesting inhibitory effects. Within the activation-inhibition framework, strong competitors were expected to yield more inhibition than weak competitors, and competition should be observed until relatively late in the target selection process. Furthermore, strong competition should be prone to errors, and erroneous selection of a higher-frequency neighbour as the target should also occur relatively late in the target selection process.

Method

Participants

A total of 121 Binghamton University students participated in the study. A large number of participants were recruited to obtain sufficient observations for analysis of errors. All participants were native speakers of English, had normal or corrected-to-normal vision with contacts, and were naïve regarding the purpose of the experiment. Participants provided written informed consent prior to participating in the study, which was approved by the University's Human Subjects Research Review Committee.

Apparatus

Stimuli were displayed on an Iiyama Vision Master Pro 514 CRT monitor with a resolution of 1,024 × 768 pixels. Target words were presented in Size 24 Courier New font, and all items in the competitor array were presented in Size 24 Times New Roman font so that selection of the target from the four forced choice alternatives could not be based on a low-level matching of graphemic features. Head position was fixed at a distance of 85 cm, and an Eye-link 1000 tracker was used to monitor eye movements at a rate of 1,000 Hz and a tracking error of 0.1° or less.

Materials

A set of 40 orthographic neighbour pairs was compiled that differed significantly in SUBTL frequency ($t = 4.66$, $df = 39$, $p < .001$, pairwise) but did not differ in number of letters, number of syllables, orthographic neighbourhood size, concreteness, or regularity (all ps > .3, pairwise). This list of stimuli was constructed using the UNION database (which merged SUBTL and Bartlett et al., 2009; Brysbaert & New, 2009; the CMU Pronunciation Dictionary Armstrong, 2013), Brysbaert et al.'s (2014) concreteness norms, and N-Watch (Davis, 2005). Each of the neighbours was also matched to an orthographic onset competitor (sharing at least the first two letters) and an orthographically unrelated control (all ps > .4, pairwise).[1] Target word properties are available in Table 1.

One member of each orthographic neighbour pair appeared as the target word in a trial. The visual array that followed presentation of the target consisted of the target

Table 1. Target word properties.

	Word type					
	LF word	LF onset	LF control	HF word	HF onset	HF control
Letters	4.95	4.95/4.95	4.95/4.95	4.95	4.95/4.95	4.95/4.95
Syllables	1.25	1.30/1.28	1.25/1.23	1.23	1.28/1.28	1.23/1.25
SUBTL frequency	11.66	11.67/55.66	11.56/68.98	69.92	65.53/12.98	69.24/11.56
N	4.73	4.70/4.83	4.83/4.80	4.83	4.73/4.55	4.88/4.83
Concreteness	3.95	4.07/4.08	4.00/3.98	4.02	3.91/4.06	3.97/4.00

Note. For the "Onset" and "Control" columns, values for the target matched and neighbour matched competitors[1] are both presented, separated by "/." HF: higher-frequency; LF: lower-frequency.

word (e.g., *spell*), its neighbour (e.g., *shell*), the target word's orthographic onset competitor (e.g., *speed*), and an orthographically unrelated control (e.g., *trade*), each positioned in a different quadrant of the screen. Five catch trials, in which the target word was not present in the array, were interspersed with the experimental trials. These trials were preceded by 15 practice trials, which had word arrays comparable to the targets but with orthographic neighbour pairs that were not used in the experimental trials. Five of these practice trials were also catch trials, in which the target was not present in the array.

Procedure

Participants were calibrated using a 9-point calibration procedure prior to beginning the 60 trials, and were recalibrated as needed throughout the experiment. At the beginning of each trial, a fixation cross appeared in the centre of each screen for 1,000 ms. Following the fixation cross, the target word appeared at the centre of the screen[2] for 40, 45, or 50 ms and then was replaced by a 50 to 60 ms visual mask[3] consisting of a row of *x*'s equal to the length of the target. Next, the forced choice visual array appeared, and the mask was replaced with a fixation cross, signalling that the participant should search for the target in the array. Participants were instructed to click on the target as quickly and accurately as possible when it was present, and to make no response if either the target was not present or if they did not know what the target word was. Participants were informed ahead of time that the target would not be present in a small subset of trials. These catch trials were included early on so that participants would focus attention on the target location and to discourage target guessing. The trial ended when the participant either made a response by clicking one of the words on the screen or failed to make a response within 3,000 ms.

Trial order was randomised within practice and experimental blocks such that practice trials (15) and the experimental and catch trials (45) appeared in different random orders for each participant. The position of the different types of words in the visual array (target, neighbour, onset competitor, and control) was counterbalanced such that each of the word types was equally likely to appear in each quadrant of the screen. The lower-frequency member of the neighbour pair was the target word in half of the trials, and the higher-frequency member was the target in the other half. The frequency of the target used was counterbalanced across two lists, such that both members of each neighbour pair appeared as the target exactly once across lists. Participants would view just one member of the neighbour pair as target, for instance, when *spell* was a target on one trial, the participant would not view *shell* as target on another trial. List assignment was counterbalanced across participants such that each member of each orthographic neighbour pair was seen by roughly half of participants.

Data selection and analysis

Due to a coding error, nine items were repeated for a subset of participants (N=41), so the repeated items were removed from the data set prior to analysis (7.6% of trials). Furthermore, trials in which participants failed to follow instructions or when tracking was lost were removed prior to analysis (65% of trials). After the above exclusion, 4,442 experimental trials were left, and this set of data was used for the analyses of target recognition errors. For analyses of first saccade latency and reaction time (RT), trials in which participants failed to make a response within 3,000 ms and trials with a first saccade latency less than 80 ms or greater than 1,000 ms were not included (10.4% of trials). Therefore, first saccade latency and RT measures were based on 3,978 trials. Analyses of fixations were limited to trials in which the correct target word was selected (83.9% of trials; i.e., 3,725 trials). Data were analysed with linear mixed models (LMMs) and generalised LMMs (GLMMs) which were implemented in R (R Core Team, 2015) using the lme4 library (Bates et al., 2015). Significance values were calculated using the lmerTest package (Kuznetsova et al., 2015), and figures were rendered using ggplot2 (Wickham, 2009).

Response accuracy, first saccade latency, and RT analyses. The numeric measures were analysed using LMMs with target frequency as a predictor and participant within target frequency condition as random effects; response accuracies were analysed using a corresponding GLMM. First saccade latency was defined as how long it took the participant to execute a saccade following the onset of the word array, and RT was defined as how long it took the participant to make a response in trials in which the target was selected. As RT and first saccade latency exhibited skewing towards longer durations, a log transformation was applied to the data prior to analysis.

Divergence analysis. To allow for analysis of competitor fixations, a 200 × 200 pixel interest area was defined around each word in the array to determine which, if any, of the four words was fixated at a given time during the trial. Analyses of fixations were separated into interest areas corresponding to the target, neighbour, onset competitor, and orthographically unrelated control. Comparisons between fixation proportions to these interest areas were conducted using divergence analysis (Dink & Ferguson, 2015). A primary advantage of this analysis technique is that it allows for identification of specific time windows during which two fixation curves differ. Furthermore, as divergence analysis is a permutation-based, nonparametric approach, it controls for multiple comparisons and is robust to non-normality.

Divergence analysis was conducted following the steps described by Dink and Ferguson (2015). First, fixation proportions to each interest area (target, neighbour, onset competitor, and unrelated distractor) were calculated for 25 ms time-bins from 250 to 1,500 ms postarray onset. For each comparison of interest (e.g., target vs. neighbour), a pairwise *t* value was calculated for each time-bin, and adjacent time-bins that exceeded a *t* threshold were grouped together into clusters. A *t*-sum value was then calculated for each cluster by summing the *t* values of each time-bin in the cluster. For example, if the *t* values for three adjacent time-bins—250 to 275 ms, 275 to 300 ms, and 300 to 325 ms—were 3, 5, and 4, respectively, then the cluster would be 250 to 325 ms, and the *t*-sum for that cluster would equal 12.

To determine whether these *t*-sums were likely to have occurred by chance, the null distribution of *t*-sums was identified by conducting 10,000 permutation tests. For each test, the data were randomly shuffled within participants, and clusters were identified following the procedure described above. The largest *t*-sum for each permutation test was stored, and *p* values were calculated by comparing the *t*-sums of the nonpermuted data set to the resulting distribution. Analyses were implemented in R using the eyetrackingR package (Dink & Ferguson, 2015).

Results

Response accuracy, first saccade latency, RT, and the time course of forced choice viewing

An examination of catch trials showed that participants made a correct response, that is, they did not click on one of the four forced choice alternatives, on 90.2% of the catch trials indicating that attention was focused at the target location at the beginning of these trials. During experimental trials, which required responding, participants moved the eyes to at least one interest area on virtually all trials (N=4,441), and the location of an interest area did not influence the frequency with which it was selected for viewing after the offset of the briefly presented central target word (χ^2=6.152, df=3, p > .1), indicating that response accuracy was not compromised by spatial preferences. As expected, this accuracy was lower (−8.2%) when the briefly presented target was the lower- rather than higher-frequency member of the critical word pair (z=−6.45, b=−.68, p < .001). The corresponding condition means and standard errors are shown in Table 2 together with the corresponding statistics for saccade latency and RT.

Table 2. Saccade latency, reaction time, and accuracy means.

	Target type	
	Higher-frequency	Lower-frequency
Saccade latency (ms)	371.32 (86.62)	376.01 (92.15)
Reaction time (ms)	1,741.80 (403.31)	1,809.57 (421.43)
Accuracy	0.88 (0.32)	0.79 (0.40)

Note. Standard deviations are listed in parentheses.

A qualitative examination of error frequencies showed that nonresponding was the most common type of error (N=451 experimental trials), again indicating that the recognition of a briefly presented target word was relatively difficult. When an erroneous choice was made, it involved primarily the selection of a neighbour in lieu of the target (N=207). Instances in which the onset competitor was selected were relatively rare (N=51), and the control word was selected on just six trials.

First saccade latency did not differ as a function of target frequency (t=1.56, b=0.01, p=.12), but there was a reliable difference in RT (t=5.72, b=.04, p < .001), such that participants were slower to respond on trials with lower-frequency targets (−68 ms). First saccade latency and RT were also significantly positively correlated (r=.50, t=6.31, df=119, p < .001), suggesting that individuals who responded more quickly were also faster to initiate saccades at the beginning of the trial.

Divergence analysis

Correct trials. Array fixation proportions are presented for targets and competitors in Figure 1. Analyses of fixation proportion were restricted to trials in which the target was correctly identified (see Error Trials for consideration of trials in which the target was not selected). To examine whether competition was evident even when the target was correctly identified, and whether neighbours served as especially strong competitors, inferential comparisons

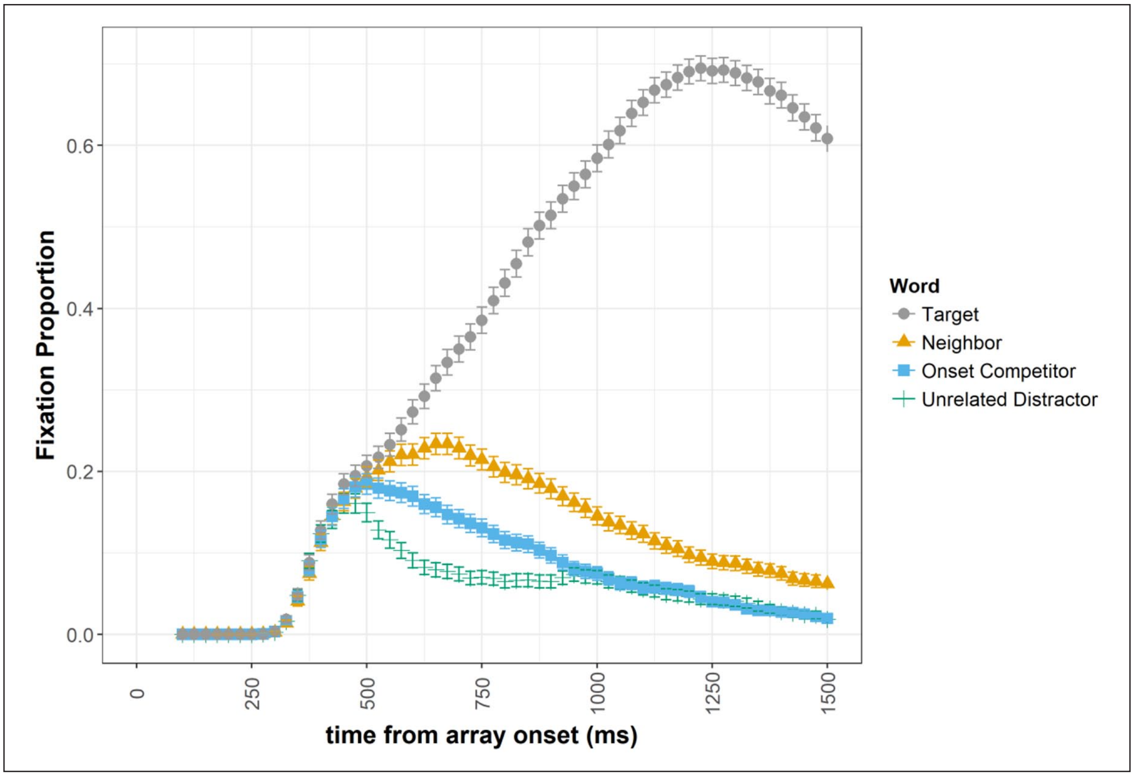

Figure 1. Fixation proportions for targets, neighbours, onset competitors, and unrelated distractors.
Note. Error bars are 95% confidence intervals.

were conducted between the neighbours and onset competitors (triangles and squares) and also between the average of these competitors and the unrelated distractors (crosses). The t threshold was based on a within-subjects two-tailed test at alpha = .05, so the critical t value was 1.98.

The fixation proportions depicted in Figure 1 suggest that participants preferentially viewed the target's neighbour over the onset competitor beginning just after 500 ms post-array onset. In line with these findings, a cluster was identified, extending from 525 ms to the end of the analysis window, *t-sum* = 246.63, and this cluster fell well outside the null distribution identified by permutation (M = 0.08, 95% confidence interval [CI] = [−19.61, 18.98], $p < .001$). For the comparison between competitors sharing orthographic overlap with the target (neighbours and onset competitors) and orthographically unrelated distractors, a significant cluster was also identified extending from 475 ms postarray onset to the end of the analysis window, *t-sum* = 336.46, $p < .001$. One other identified cluster fell within the null distribution (M = 0.04, 95% CI = [−18.49, 18.66], p = .47). Together, these findings indicate that orthographically related competitors were preferentially viewed over unrelated distractors, and that neighbours served as stronger competitors than onset competitors.

Another question of interest was whether competition was modulated by neighbour frequency. Figure 2 shows fixation proportions to targets and competitors as a function of target frequency. To determine whether neighbour frequency influenced the time course of competition, comparisons were carried out between the target and neighbour for higher- and lower-frequency targets separately. For the higher-frequency target versus neighbour comparison, a significant cluster was identified from 525 ms postarray onset to the end of the analysis window, *t-sum* = 1,008.76, $p < .001$. All other identified clusters fell within the null distribution (M = −0.32, 95% CI = [−27.79, 22.86], $ps > .5$). In comparison, the neighbour diverged from the target somewhat later for the lower-frequency target trials. Specifically, a significant cluster was identified beginning 750 ms postarray onset, *t-sum* = 269.38, $p < .001$, and one other cluster fell within the null distribution (M = −0.10, 95% CI = [−26.46, 25.89], p = .11).

Figure 3 shows fixation proportions to the neighbour as a function of target frequency, and inspection of this figure suggests that higher-frequency neighbours of lower-frequency targets were preferentially viewed over lower-frequency neighbours over much of the analysis window. To confirm this, fixation proportions between higher-frequency neighbours and lower-frequency neighbours were compared directly, and the outcome of this analysis is depicted in Figure 4. Two clusters were identified from 500 to 850 ms postarray onset, *t-sum* = −48.21, p = .01, and from 875 ms to the end of the analysis window,

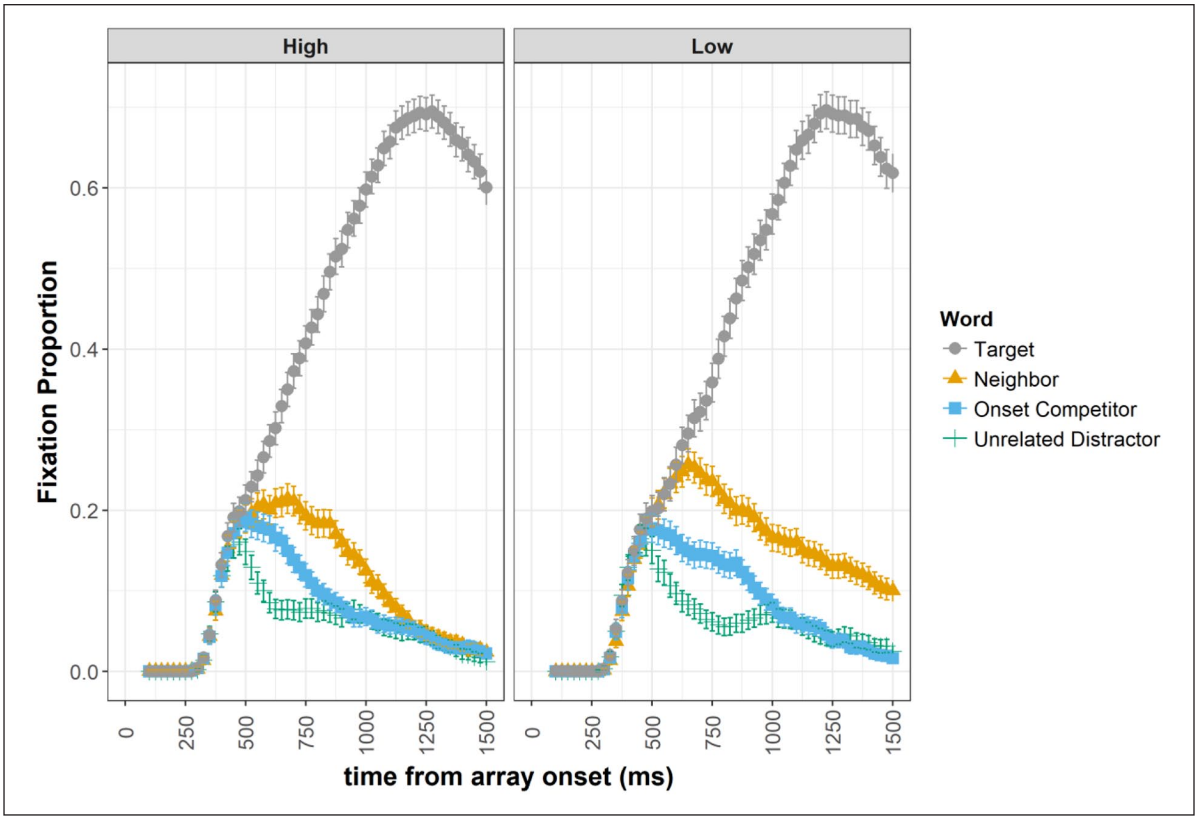

Figure 2. Fixation proportions for targets, neighbours, onset competitors, and unrelated distractors as a function of target frequency.
Note. Error bars are 95% confidence intervals.

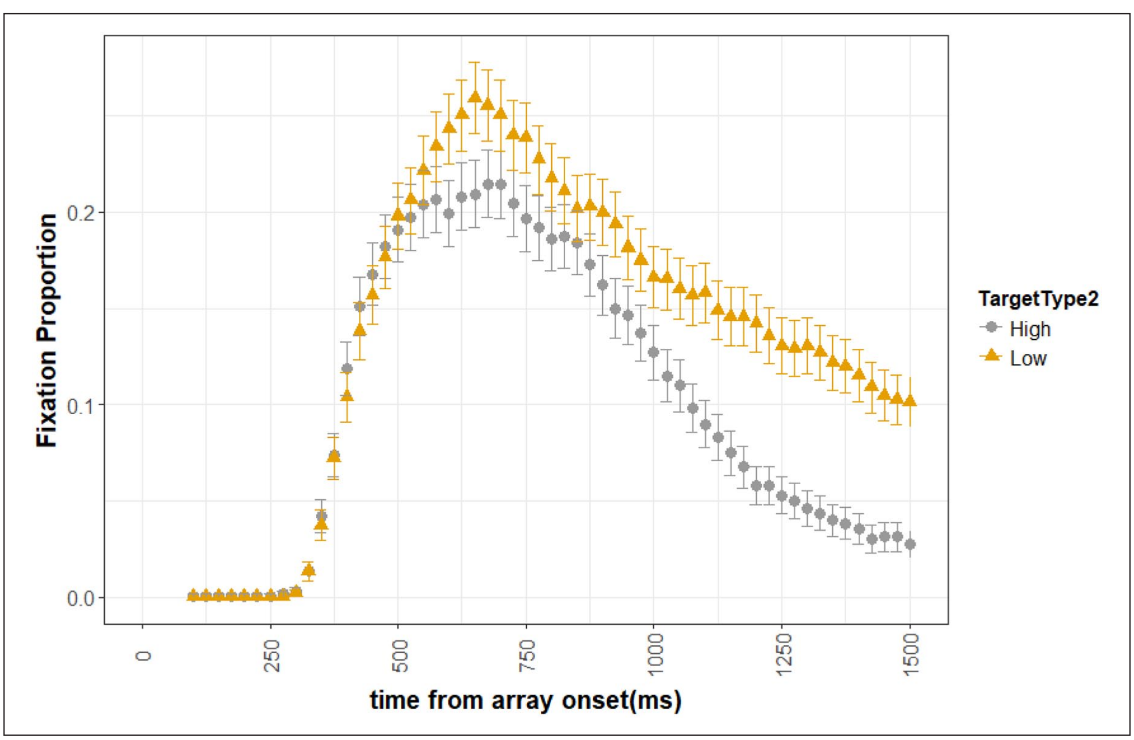

Figure 3. Fixation proportions for neighbours as a function of target frequency.
Note. Error bars are 95% confidence intervals.

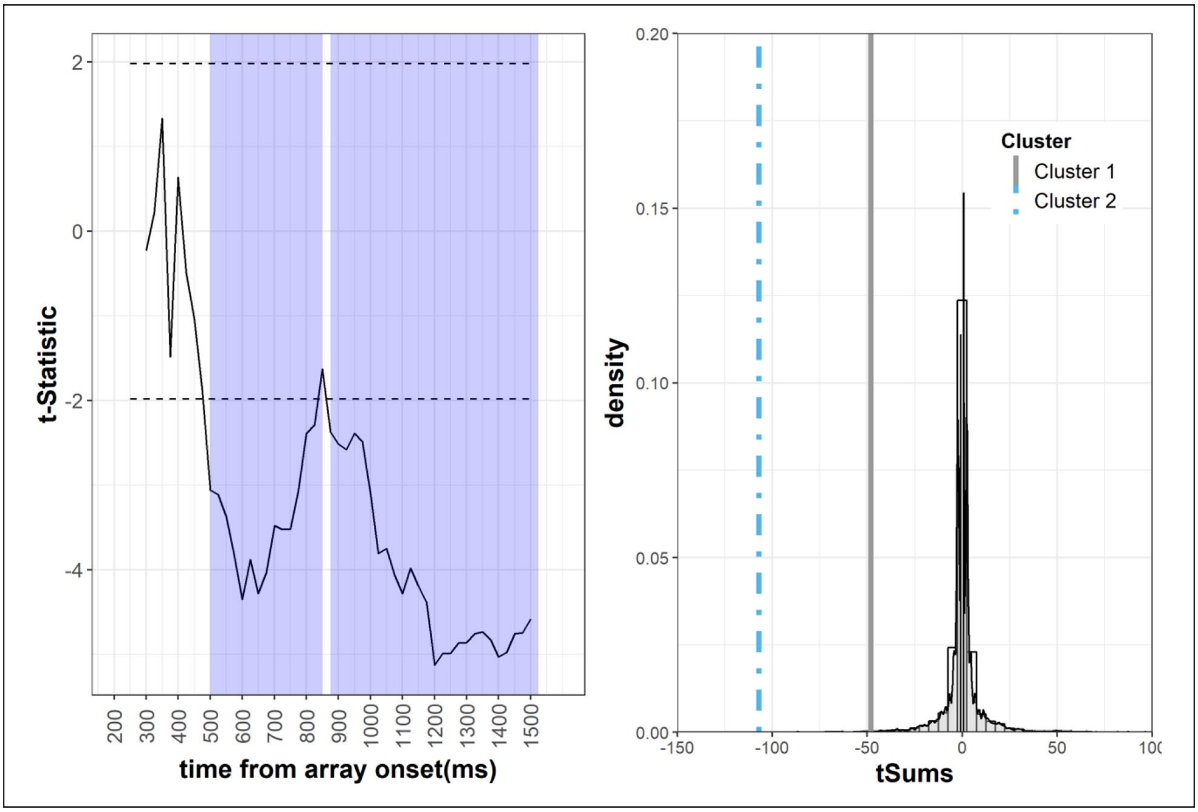

Figure 4. The outcome of divergence analysis between neighbours of high- and low-frequency targets.
Note. The left panel shows the *t* statistics for the high versus low comparison over time. The horizontal lines represent the threshold for significance based on a two-tailed test at alpha = .05, and time-bins that exceeded this threshold are shaded. The right panel shows the distribution of 10,000 permuted *t*-sums, how the *t*-sums of identified clusters compare to this distribution, and that both clusters fell outside of the null distribution. Clusters are labelled in temporal order.

t-sum = −107.15, $p < .001$. Both of these clusters fell outside the null distribution of *t*-sums identified by permutation ($M = -0.05$, 95% CI = [−20.22, 19.82]), and no other clusters were identified. Taken together, these results indicate that higher-frequency neighbours were preferentially viewed over lower-frequency neighbours beginning around 500 ms after array onset, and that viewing preference for the neighbours diverged from the target later when the neighbour was higher in frequency than the target (lower-frequency target trials).

Error trials. A motivating factor for conducting this experiment was to explore the time course of neighbour competition when an error is made, specifically when the participant selects the incorrect word from the array. As reported before, when an incorrect forced choice alternative was selected instead of the target, the neighbour was chosen on the vast majority of trials. Given this information, it does not appear to be the case that errors in which a competitor other than the target word was selected simply represent random misclicks; instead, it can be characterised as a misreading of the target word as another word in the array.

To examine the time course of neighbour competition when an error was made, divergence analyses were conducted on fixation proportions for trials in which the neighbour was selected in lieu of the target. As 89 of the participants made this error, the *t* threshold for this comparison was 1.99. Figure 5 shows fixation proportions to the target (circles) and neighbour (triangles) as a function of target frequency for these trials (67 trials with higher-frequency targets, and 140 trials with lower-frequency targets). As is clear from the figure, when the neighbour was erroneously selected, there was a corresponding viewing preference for the neighbour over the target. In line with this, when target and neighbour fixation proportions were compared, a cluster indicating viewing preference for the neighbour was identified from 500 ms postarray onset until the end of the analysis window, *t-sum* = −206.00, $p < .001$, which fell outside the null distribution identified by permutation ($M = -0.02$, 95% CI = [−19.91, 19.73]).

In addition, inferential comparisons between targets and neighbours were conducted separately for higher- and lower-frequency target error trials to examine the effect of target frequency on error making. As 45 participants misread higher-frequency targets as their neighbours and 76

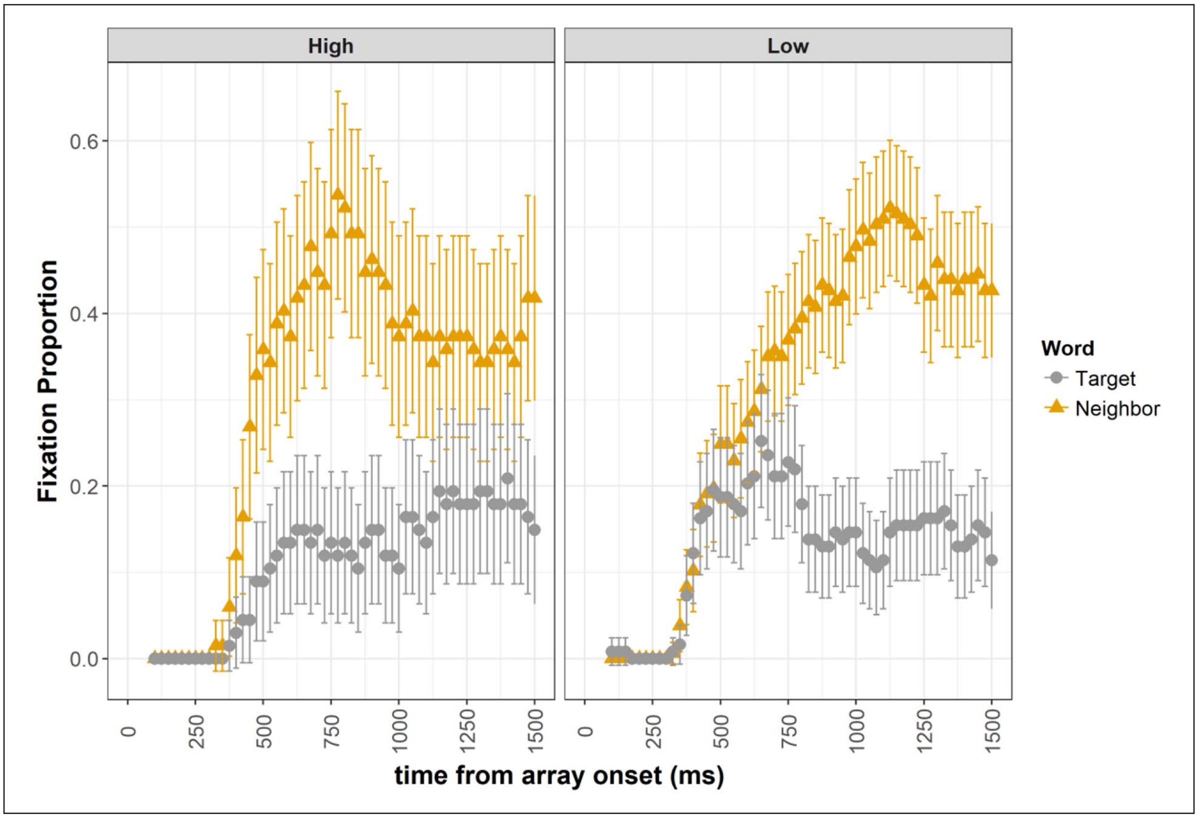

Figure 5. Fixation proportions for targets and neighbours as a function of target type for trials in which the target was misidentified as its neighbour.

participants misread lower-frequency targets, the critical t values for these comparisons were 2.02 and 1.99, respectively. In error trials with higher-frequency targets, a significant cluster was identified from 425 to 1,175 ms postarray onset, $t\text{-}sum = -96.86$, $p < .001$. For trials with lower-frequency targets, preference for the neighbour emerged comparatively late, and a reliable cluster was identified from 675 ms to the end of the analysis window, $t\text{-}sum = -184.99$, $p < .001$. All other identified clusters fell within the null distribution for trials with higher-frequency targets ($M = 0.14$, 95% CI = [−20.66, 20.94], $ps > .2$) and lower-frequency targets ($M = 0.01$, 95% CI = [−20.18, 20.37], $p > .7$). Overall, neighbours were preferentially viewed over targets when they were misread, and this preference began and ended earlier for trials with higher- than with lower-frequency targets.

Discussion

This study was conducted to examine the generality of competition based on orthographic similarity during visual word recognition, and to know whether existing findings may be attributed to the occasional making of errors. To accomplish this, we used a modified version of the VWP with visual rather than spoken words. Following brief presentation of visual target word, participants were asked to select this word from a printed forced choice array containing the target, its higher- or lower-frequency orthographic neighbour, an orthographic onset competitor, and an orthographically unrelated distractor. Behavioural responses and viewing preferences were examined to determine whether competition was observed when the target was identified correctly, and whether errors arose from perceptual or competition-based processes.

Altogether, the current findings suggest that competition occurs generally and is sensitive to both orthographic similarity and relative frequency. When analyses were restricted to trials in which the target was correctly identified, neighbours were preferentially viewed over orthographic onset competitors, and both orthographically related competitors were preferentially viewed over unrelated distractors. This influence of orthographic similarity on competition was also evident in the pattern of error making, as the majority of errors reflected misreading of the target as its neighbour, and the orthographically unrelated distractor was almost never selected in lieu of the target.

Relative frequency of the target and its neighbour also influenced competition. Participants were slower and less accurate at selecting the target when the forced choice array contained its higher-frequency neighbour. Furthermore, when the target was correctly identified, higher-frequency neighbours served as stronger competitors than lower-frequency neighbours; in comparison with lower-frequency neighbours, higher-frequency neighbours received a higher fixation proportion. Divergence analyses further indicated that the target was chosen over the neighbour at a later point in time. This is in line with other findings indicating that the presence of higher-frequency neighbours is inhibitory (Huntsman & Lima, 1996; Paterson et al., 2009; Perea & Pollatsek, 1998; Sears et al., 2006, Experiment 1A), and that it may even lead to a recognition error (Gregg & Inhoff, 2016; Potter et al., 1993; Slattery, 2009; Warrington et al., 2016).

Error making was similarly influenced by orthographic similarity and relative frequency. As mentioned above, errors occurred as a function of orthographic similarity with the target and were more likely when a higher-frequency neighbour was present in the array. Although the influence of orthographic similarity is consistent with both perceptual and competition-based accounts of error making, the influence of relative frequency suggests that errors arise at least in part due to ongoing lexical competition between the target and activated competitors. That is, given that higher- and lower-frequency neighbours are equally perceptually similar to the target (differing by a single letter), the influence of relative frequency suggests that higher-frequency neighbours were more strongly activated during word identification, and this increased the likelihood that the word was misread.

To further examine the extent to which misreading resulted from perceptual or competition-based processes, fixation proportion was examined for trials in which the target was misread as its neighbour. For these trials, viewing preference for the neighbour occurred later in the time course for words with higher-frequency compared with lower-frequency neighbours. Although these results are preliminary, this suggests that there may be two kinds of errors that differ in their time course: early, perceptual errors, in which the target is misperceived as another word from an early stage; and misidentifications, in which the incorrect word is selected from among lexical candidates later in the time course of word identification.

Time course of competition

Although effects of neighbour frequency are generally inhibitory in other languages, in English, the effect is somewhat more inconsistent, with only some studies showing inhibitory effects, and others showing null or facilitatory effects. This inconsistent result pattern may result from shifts in the relative contribution of lexical-level inhibition versus sublexical facilitation from a word's orthographic neighbourhood (McClelland & Rumelhart, 1981), and thus even facilitatory effects do not necessarily contradict the occurrence of inhibition between activated lexical candidates. In other words, research on orthographic neighbourhood can be described in terms of a broader framework in which early facilitative effects versus late inhibitory effects can be attributed to different stages of processing during word identification (see Perea & Rosa, 2000, for a review). That is, the presence of neighbours may be facilitative during an assessment of whether a target is a word (as occurs in a lexical decision task) but inhibitory when selecting among competing lexical candidates (note that this is consistent with the architecture of the E-Z Reader model; see Reichle et al., 2006, for a review of the E-Z reader model). Broadly, this is consistent with lateral inhibition accounts of lexical selection in existing models (e.g., Davis, 2010; McClelland & Rumelhart, 1981; Paap et al., 1982; Whitney, 2001).

Given that competition effects were captured by a task that began after target offset, within this framework only inhibitory effects of neighbour frequency would be expected to emerge. In line with this, the earliest measure, saccade latency, did not differ as a function of target frequency, and thus did not show evidence of inhibition or facilitation. However, the effect of neighbour frequency was decidedly inhibitory later in the time course, as the presence of a higher-frequency neighbour increased RT, decreased accuracy, and led to greater and more sustained competition until at least 1,500 ms post-target offset.

The relatively long target-neighbour competition is difficult to reconcile with models of word recognition during reading, such as the E-Z Reader model (Reichle et al., 2006), which assume that the recognition of a word is concluded relatively quickly so that attention can be focused on the next word. Although it is the case that the target was likely removed from the screen before identification could be completed in some cases, it is nonetheless surprising that competition effects would persist as long as 1,500 ms after target offset, considering that fixation durations during reading are typically 200 to 250 ms (Rayner, 1998). In this study, in contrast, competition during word identification does not appear to be necessarily resolved during first pass of a word, as it is evident even after the target word is no longer present on the screen. This is consistent with the findings of previous studies suggesting that readers may maintain uncertainty about, and even revise, word identity after first pass of the word (Gregg & Inhoff, 2016; Levy et al., 2009).

Furthermore, this suggests that inhibitory neighbour frequency effects, which have been attributed to word identification errors, could also, in part, reflect ongoing competition after first pass (Acha & Perea, 2008; Gregg & Inhoff, 2016; Pollatsek et al., 1999; Slattery, 2009; Warrington et al., 2016). That is, late inhibitory effects may have arisen either due to corrective responding due to failure to integrate an erroneous word identity into subsequent context, or efforts to resolve lingering uncertainty about word identity between multiple activated lexical candidates, which would be particularly likely for words with higher-frequency neighbours.

The proposed lexical competition account is predicated on the assumption that the observed neighbour frequency effects reflect general target recognition processes rather than episode-specific influences of the visual array. Specifically, it could be argued that the relative activation strength of the neighbour words depends on the match between the memory trace of the target and the words that are viewed when the forced choice array is presented. Given that the target was presented briefly, this memory trace is likely to be somewhat degraded and, since the target and neighbour shared all but one letter, the degraded memory trace will match the target and neighbour more strongly than the onset competitor and the unrelated alternative. Assuming that a higher-frequency neighbour among the forced choice alternatives will be activated more strongly than a lower-frequency neighbour when the memory trace more closely matches the other member of the neighbour pair, the selection of a lower-frequency target should be relatively difficult, that is, take more time and be more error-prone. That is, the recognition of the higher-frequency neighbour on the forced choice display dominates target recognition processes, and this may occur even when the memory trace for the lower-frequency target correctly coded the critical discriminating letter. In contrast to this, neighbour errors to higher-frequency targets will be limited to trials on which the target trace is more degraded, particularly for the critical letter, so that the lower-frequency neighbour is selected among the four forced choice alternatives before the higher-frequency target is viewed.[4]

To examine this theoretical alternative, the examination of response accuracy was expanded so it included the type of forced choice viewing. According to the alternative, the recognition of a lower-frequency target should be less susceptible to error when the initially viewed forced choice is the same word as the target, and the recognition of a higher-frequency target should be more error-prone, when lower-frequency neighbour is viewed immediately after the target's presentation.

In a first step, we determined the type of forced choice word that was fixated after the offset of the target's presentation. It showed a slightly lower selection rate for control words (23.4%) than for onset competitors (25.6%), lower-frequency members of neighbour pairs (25.1%), and higher-frequency members of pairs (25.8%). After this, we examined response accuracy on experimental trials as a function of target frequency and the forced choice alternative that was viewed first (Table 3). Sliding contrasts were applied to the type of forced choice word (target vs. neighbour, neighbour vs. onset competitor, onset competitor vs. control), and the random factor structure included participants and their slopes for target frequency. The results showed a robust effect of target frequency, but no significant effect for the initially fixated forced choice type. Importantly, none of the interactions between target frequency and the type of fixated forced alternative approached significance, all $ps > .154$. These findings disagree with the alternative, as the viewing of the forced choice display did not determine the accuracy of target recognition.

Table 3. Accuracy as a function of first force choice viewing.

Interest area	Target type	
	Higher-frequency	Lower-frequency
Target	0.92 (0.01)	0.82 (0.02)
Neighbour	0.90 (0.01)	0.83 (0.02)
Onset competitor	0.89 (0.01)	0.83 (0.02)
Unrelated distractor	0.91 (0.01)	0.84 (0.02)

Note. Standard deviations are listed in parentheses.

Although this study provides evidence that both error making and ongoing competition play a role in word recognition, further research is needed to characterise the contributions of enduring lexical competition and error making to neighbour frequency effects during normal silent reading.

Time course of errors

A motivating question was to determine whether errors originate from issues early in processing, late in processing, or both. Analysis of error trials suggested that there may be two distinct kinds of errors: misperceptions and misidentifications. Specifically, errors in which a lower-frequency neighbour was selected in lieu of a higher-frequency target appeared to manifest earlier in the time course than trials in which a lower-frequency target was misread as its higher-frequency neighbour. Together, these observations suggest that errors can either become evident early during the time course due to incorrect initial encoding of the word's orthographic information (misperception) or later failure to correctly select from among the target word's competitors (misidentification). Furthermore, while misperceptions may occur regardless of the relative frequency of the target and its neighbour, misidentifications primarily happen when the neighbour is higher in frequency, and consequently, they are seldom observed for higher-frequency targets. This suggests that errors can originate both early and late in processing, and that errors may be broken down into two subcategories. Furthermore, of these two categories, misidentification is more likely, but only when the target word's neighbour is higher in frequency. When this is not the case, the small number of observed errors seems to reflect misperceptions rather than misidentifications.

Given the possibility that there may be two common forms of errors, why would one type of error be more prevalent for a higher-frequency target while another would occur more often for a lower-frequency target? One possibility concerns the nature of competition among lexical candidates during word identification. As has been observed in previous experiments (e.g., Gregg & Inhoff, 2016; Pollatsek et al., 1999), the nature of neighbour competition and misreading seems to be directional by frequency, such that a higher-frequency neighbour is a strong competitor while lower-frequency neighbours act as weaker competitors. In an abstract sense, one can think of selection among lexical candidates as being weighted by various types of evidence (e.g., orthographic information, relative frequency of alternatives, context) and by the fluency of processing (accrual of activation). If the target is the lower-frequency member of the neighbour pair, the competition between the different types of evidence may be relatively strong because the accrual of activation is relatively slow for both competing members of the neighbour pair, suggesting that the briefly presented target could have been the lower-frequency member, while the relative frequency of alternatives suggests the higher-frequency member. It is plausible that resolution of this conflict takes time and is relatively error-prone.

However, slow errors would be less likely for a higher-frequency target, as the evidence supporting the lower-frequency neighbour as the target is weaker (it shares orthographic information, but it is generally less likely to be present in text). In this case, an erroneous choice may occur when early perceptual processing is compromised in some form. It could also occur when the reader is confronted with an unexpected processing outcome. That is, recognition of the briefly presented target failed, even though its processing was relatively fluent. When this occurs, the reader may conclude that the target was the lower rather than the higher-frequency member of the neighbour pair.

However, this kind of error would be substantially less likely for a higher-frequency target, as the evidence supporting the lower-frequency neighbour as the target is weaker (it shares orthographic information, but it is generally less likely to be present in text). In this case, observed errors with higher-frequency targets may occur primarily due to early perceptual issues in which the visual information about the word is somehow misrepresented. Although these errors would also be possible with lower-frequency targets, later time course errors would be generally more likely and disguise evidence in the data for the earlier, perceptual errors.

Altogether, errors most often involved selection of the neighbour from the array, and errors were more likely to occur when the target was lower-frequency. Analysis of fixation proportion during error trials suggests that lower-frequency target errors may primarily occur during later selection from lexical candidates, while higher-frequency target errors may reflect a rarer perceptual failure early during the time course of word identification.

Using the VWP to study visual word identification

Given its sensitivity to ongoing competition effects, the VWP has been extensively used to study spoken word processing for the last 20 years. The results of this study, in addition to those of Meyer and Federmeier (2008), suggest that this paradigm can be successfully adapted to study ongoing competition following visual word identification, as well. Although this paradigm is novel and deviates substantially from normal silent reading, resulting viewing proportions coalesce neatly with behavioural measures (RT and accuracy) into an internally consistent account of competition and error making that is also entirely plausible within the context of existing literature.

One potential concern is that given its similarity to visual search, the present adaptation of the VWP may primarily reflect visual similarity between the target and members of the array, rather than overlap between the members of the array and activated lexical representations. The results of this study and Meyer and Federmeier's (2008) study rule out this possibility. Specifically, Meyer and Federmeier showed evidence for preferential viewing of array members as a function of semantic rather than orthographic (visual) overlap with the target. Furthermore, in this study, neighbour frequency influenced ongoing competition in addition to orthographic similarity. As both higher- and lower-frequency neighbours were equally and perceptually similar to the target, similarity-based visual search of the array cannot account for this finding. Altogether, the current paradigm provides a novel and promising way to examine ongoing competition effects following visual word identification, and it is sensitive to the overlap between members of the array and activated lexical representations.

Conclusion

The findings of this study suggest that competition occurs generally, not just when words are misidentified. When the target was correctly identified, competition occurred as a function of orthographic similarity with the target, and it was modulated by target frequency, such that higher-frequency neighbours served as stronger competitors than lower-frequency neighbours. The contribution of orthographic similarity and neighbour frequency to ongoing competition was also evident in the pattern of errors, as participants were most likely to misidentify targets as their higher-frequency neighbours. Furthermore, examination of error trials provides preliminary evidence that errors may occur either as a function of early perceptual failure (misperception) or due to incorrect resolution of ongoing competition (misidentification). Overall, this paradigm provides a sensitive means to observe whether neighbours compete for activation and the extent to which this activation is strong enough to influence word identification. More generally, it provides a promising mechanism for studying the activation of lexical representations during visual word identification. Future work should consider the potential contribution of errors to lexical competition and could accomplish this by applying this adaptation of the VWP.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a grant from the National Natural Science Foundation of China (31970992), and a grant from National Social Science Foundation of China (20&ZD296). This work was also jointly funded by the National Natural Science Foundation of China (NSFC) and the German Research Foundation (DFG) in Project Crossmodal Learning, NSFC 62061136001/DFG TRR-169.

ORCID iD

Xingshan Li https://orcid.org/0000-0002-0995-4631

Notes

1. In a subset of the data ($N=39$), the frequency of the onset competitor and of the unrelated distractor was matched to the target, and for the remainder of the data ($N=82$), onset competitors and distractors were matched in frequency to the target's neighbour within a display. As the relative frequency of these competitors did not alter the reported effect pattern, data are presented collapsed across these conditions.
2. Pilot studies were used to find an optimal target presentation duration. We started with a duration of 35 ms (with $N=10$ participants) which yielded a relatively low response accuracy of 65%, suggesting that target selection in this condition was largely based on guessing. The accuracy increased to 71%, 84%, and 87% with 40 ms ($N=11$), 45 ms ($N=15$), and 50 ms target presentation durations. Because errors in the two shorter presentation condition involved more often nonresponding, we selected the 50 ms target presentation duration for the experiment ($N=95$).

   A supplementary analysis of response accuracy was performed to determine whether increases in target durations would change the viewing of the four forced choice alternatives. For this, we examined response accuracy as a function of target duration and the type of word that was initially selected for viewing when the four forced choice alternatives appeared (target, neighbour, onset competitor, and control word). The statistical model included target duration, the initially selected type of word (coded as sliding contrasts: target vs. neighbour, neighbours vs. onset competitor, and onset competitor vs. control), and the interactions of duration with each sliding contrast as predictors, and participants within target frequency were again used as random factors. The results showed a robust effect of target duration ($b=0.0148$, $SE=0.0031$, $t=4.853$, $p<.001$), as accuracy increased with duration, but there was virtually no effect for the type of initially viewed forced choice word and no reliable interaction involving the forced choice word (all $ps>.28$), thus indicating that differences in target duration did not influence the initial viewing of the forced choice display.

   To allow for qualitative analysis of as many errors as possible, we thus compounded the data from the 40, 45, and 50 ms target durations.
3. To simplify experiment programming and analysis, the mask duration was changed between experiments so that the summed target and mask durations always equalled 100 ms.
4. We are grateful to one of the reviewers for pointing out this alternative account for our data.

References

Acha, J., & Perea, M. (2008). The effect of neighborhood frequency in reading: Evidence with transposed-letter neighbors. *Cognition*, *108*(1), 290–300. https://doi.org/10.1016/j.cognition.2008.02.006

Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language*, *35*, 775–800.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, *4*, 439–461.

Armstrong, B. (2013, May 20). *UNION: Merge of SUBTL and the CMU pronunciation Dictionary—Expanded*. http://blairarmstrong.net/tools/Union_Subtl_CMUPron/index.html

Bartlett, S., Kondrak, G., & Cherry, C. (2009). *On the syllabification of phonemes*. NAACL-HLT.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *Lme4: Linear mixed-effects models using Eigen and S4* (R Package version 1.0-12). http://CRAN.R-project.org/package=lme4

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990.

Brysbaert, M., Warriner, A., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *23*(4), 857–871. https://doi.org/10.1037/0278-7393.23.4.857

Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Verbal Behavior*, *18*(2), 225–241.

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical

neighbors. *Psychological Review, 119*(2), 417–430. https://doi.org/10.1037/a0027175
Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*, 84–107.
Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70.
Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713–758. https://doi.org/10.1037/a0019738
Davis, C. J., Perea, M., & Acha, J. (2009). Re(de)fining the orthographic neighborhood: The role of addition and deletion neighbors in lexical decision and reading. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(5), 1550–1570. https://doi.org/10.1037/a0014253
Dink, J. W., & Ferguson, B. (2015). *eyetrackingR: An R library for eye-tracking data analysis*. http://www.eyetrackingr.com
Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777–813. https://doi.org/10.1037/0033-295X.112.4.777
Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(3), 696–713. https://doi.org/10.1037/0278-7393.22.3.696
Grainger, J. (1990). Word frequency and neigborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, *29*, 228–244.
Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*(3), 518–565.
Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psychophysics*, *45*(3), 189–195. https://doi.org/10.3758/BF03210696
Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics*, *47*(2), 191–198. https://doi.org/10.3758/BF03205983
Gregg, J., & Inhoff, A. W. (2016). Misperception of orthographic neighbors during silent and oral reading. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 799–820. https://doi.org/10.1037/xhp0000193
Heuven, W. J. B., Van Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *483*, 458–483.
Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003
Huntsman, L. A., & Lima, S. D. (1996). Orthographic neighborhood structure and lexical access. *Journal of Psycholinguistic Research*, *25*, 417–429. https://doi.org/10.1007/BF01727000
Huntsman, L. A., & Lima, S. D. (2002). Orthographic neighbors and visual word recognition. *Journal of Psycholinguistic Research*, *31*(3), 289–306. https://doi.org/10.1023/A:1015544213366
Johnson, R. L. (2009). The quiet clam is quite calm: Transposed-letter neighborhood effects on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 943–969. https://doi.org/10.1037/a0015572
Johnson, R. L., Staub, A., & Fleri, A. M. (2012). Distributional analysis of the transposed-letter neighborhood effect on naming latency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(6), 1773–1779. https://doi.org/10.1037/a0028222
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). *lmerTest: Tests for random and fixed effects for linear mixed effect models* (lmer objects of lme4 package, R package version 2.0-3). http://CRAN.R-project.org/package=lmerTest
Levy, R., Bicknell, K., Slattery, T. J., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, *106*(50), 21086–21090. https://doi.org/10.1073/pnas.0907664106
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*(5), 375–407.
Meyer, A. M., & Federmeier, K. D. (2008). The divided visual world paradigm: Eye tracking reveals hemispheric asymmetries in lexical ambiguity resolution. *Brain Research*, *1222*, 166–183. https://doi.org/10.1016/j.brainres.2008.05.033
Paap, K. R., Johansen, L. S., Chun, E., & Vonnahme, P. (2000). Neighborhood frequency does affect performance in the Reicher task: Encoding or decision? *Journal of Experimental Psychology: Human Perception and Performance*, *26*(6), 1691–1720. https://doi.org/10.1037//0096-1523.26.6.1691
Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word-superiority effect. *Psychological Review*, *89*(5), 573–594. https://doi.org/10.1037/0033-295X.89.5.573
Pagán, A., Paterson, K. B., Blythe, H. I., & Liversedge, S. P. (2015). An inhibitory influence of transposed-letter neighbors on eye movements during reading. *Psychonomic Bulletin & Review*, *23*, 278–284. https://doi.org/10.3758/s13423-015-0869-5
Paterson, K. B., Liversedge, S. P., & Davis, C. J. (2009). Inhibitory neighbor priming effects in eye movements during reading. *Psychonomic Bulletin & Review*, *16*(1), 43–50. https://doi.org/10.3758/PBR.16.1.43
Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 767–779. http://www.ncbi.nlm.nih.gov/pubmed/9627415
Perea, M., & Rosa, E. (2000). The effects of orthographic neighborhood in reading and laboratory word identification tasks: A review. *Psicológica*, *21*, 327–340.

Pollatsek, A., Perea, M., & Binder, K. S. (1999). The effects of "neighborhood size" in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1142–1158.

Potter, M. C., Moryadas, A., Abrams, I., & Noel, A. (1993). Word perception and misperception in context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(1), 3–22.

R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. Journal of Experimental Psychology: *Human Perception and Performance, 7*(1), 167–179. https://doi.org/10.1037//0096-1523.7.1.167

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E–Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, *7*(1), 4–22. https://doi.org/10.1016/j.cogsys.2005.07.002

Sears, C. R., Campbell, C. R., & Lupker, S. J. (2006). Is there a neighborhood frequency effect in English? Evidence from reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1040–1062. https://doi.org/10.1037/0096-1523.32.4.1040

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 876–900. https://doi.org/10.1037/0096-1523.21.4.876

Slattery, T. J. (2009). Word misperception, the neighbor frequency effect, and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1969–1975. https://doi.org/10.1037/a0016894

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Warrington, K. L., White, S. J., & Paterson, K. B. (2016). Ageing and the misperception of words: Evidence from eye movements during reading. *The Quarterly Journal of Experimental Psychology*, *71*, 75–84. https://doi.org/10.1080/17470218.2016.1251471

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*(2), 221–243. https://doi.org/10.3758/BF03196158

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer.

Yao, P., Slattery, T. J., & Li, X. (2021). Sentence context modulates the neighborhood frequency effect: Evidence from eye movements while reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0001030