Orthographic and feature-level contributions to letter identification

Journal:	*Quarterly Journal of Experimental Psychology*
Manuscript ID	QJE-STD-22-076.R1
Manuscript Type:	Standard Article
Date Submitted by the Author:	18-May-2022
Complete List of Authors:	Lally, Clare; University College London, Rastle, Kathy; Royal Holloway and New Bedford College,
Keywords:	visual word recognition, reading, letter identification, visual processing, orthographic processing

DOI: 10.1177/17470218221106155

Orthographic and feature-level contributions to letter identification

Clare Lally[1,2] & Kathleen Rastle[1]

1. Royal Holloway, University of London
2. University College London

RUNNING HEAD: ORTHOGRAPHIC AND FEATURE-LEVEL CONTRIBUTIONS TO LETTER IDENTIFICATION

Correspondence address:

Clare Lally

UCL Speech, Hearing and Phonetic Sciences

Chandler House

2 Wakefield Street

London WC1N 1PF

Email: c.lally@ucl.ac.uk

DOI: 10.1177/17470218221106155

Abstract

Word recognition is facilitated by primes containing visually similar letters (*dentjst-dentist*, Marcet & Perea, 2017), suggesting that letter identities are encoded with initial uncertainty. Orthographic knowledge also guides letter identification, as readers are more accurate at identifying letters in words compared to pseudowords (Reicher, 1969; Wheeler, 1970). We investigated how higher-level orthographic knowledge and low-level visual feature analysis operate in combination during letter identification. We conducted a Reicher-Wheeler task to compare readers' ability to discriminate between visually similar and dissimilar letters across different orthographic contexts (words, pseudowords, and consonant strings). Orthographic context and visual similarity had independent effects on letter identification, and there was no interaction between these factors. The magnitude of these effects indicated that higher-level orthographic information plays a greater role than lower-level visual feature information in letter identification. We propose that readers use orthographic knowledge to refine potential letter candidates while visual feature information is accumulated. This combination of higher-level knowledge and low-level feature analysis may be essential in permitting the flexibility required to identify visual variations of the same letter (e.g. N-n) whilst maintaining enough precision to tell visually similar letters apart (e.g. n-h). These results provide new insights on the integration of visual and linguistic information and highlight the need for greater integration between models of reading and visual processing.

Keywords: *visual word recognition, reading, letter identification, visual processing, orthographic processing*

This study was pre-registered on the Open Science Framework. Pre-registration, stimuli, instructions, trial-level data, and analysis scripts are openly available (https://osf.io/p4q9u/).

DOI: 10.1177/17470218221106155

Introduction

Understanding the processes that underpin letter identification has been a long-standing goal within experimental psychology. Readers must maintain enough flexibility to recognise that *gate* and *GATE* are the same word, but also enough precision to recognise that *gate* and *gale* are not. Research shows that readers activate letter representations rapidly despite wide-ranging variability in their visual form (e.g. case and font; Bowers, et al., 1998; Hannagan et al., 2012; Kinoshita & Kaplan, 2008). However, existing literature also reveals that this flexibility extends beyond letter identity in the initial moments of visual word recognition. Masked priming paradigms demonstrate that word recognition is facilitated by prior presentation of stimuli that contain visually similar letters (*dentjst-DENTIST vs. dentgst-DENTIST,* Marcet & Perea, 2017*; docurnent–DOCUMENT* vs. *docusnent–DOCUMENT*; Marcet & Perea, 2018a), numbers (*C4BLE-cable* vs. *C9BLE-cable;* Kinoshita et al., 2013; Perea et al., 2008) or symbols (*C⊿BLE-CABLE;* Perea et al., 2008). Evidence from eye-tracking shows facilitation from visual feature similarities, shown by shorter fixation times for target words in sentences when parafoveal preview contains a pseudoword neighbour with a visually similar letter compared to a visually dissimilar letter (e.g. frjed-fried vs. frged-fried; Marcet & Perea, 2018b). ERP data also demonstrate that strings containing letter-like numbers can facilitate lexical access, as such strings evoke similar N400 semantic responses to the words they resemble (*4PPL3-APPLE;* Lien et al., 2014). Together, these findings suggest that the process of letter identification may consist of an accumulation of information about visual features.

Readers draw upon their knowledge of the writing system to support letter identification processes. For example, readers adjust prioritisation of different visual features as they gain expertise in an unfamiliar alphabet, in order to discriminate between letters (Wiley et al., 2016). Letter identification is guided by orthographic knowledge, such as knowledge of legal letter

DOI: 10.1177/17470218221106155

combinations or existing words. Consequently, the contexts in which letters appear can significantly alter readers' ability to discriminate between them. Readers identify letters more accurately when they appear in a real word compared to a pseudoword (Coch & Mitra, 2010; Grainger & Jacobs, 1994; Kezilas et al., 2016; Reicher, 1969; Wheeler, 1970). This *word superiority effect* is understood as evidence that word representations enrich letter identification processes (Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). Letter identification is also more accurate in pronounceable pseudowords (*pable*) compared to unpronounceable consonant strings (*pkwtj*) (Baron & Thurston, 1973; Carr et al., 1978). This *pseudoword superiority effect* suggests that letter identification is also guided readers' knowledge of orthotactic constraints (i.e. restrictions on how letters combine within a writing system; Kezilas et al., 2016). Thus, orthographic knowledge appears to play a key role in resolving early uncertainty around letter identity, and may reduce confusability from shared letter features. However, this line of research has not generally tested or controlled for effects of visual feature similarity.

Other work has explored whether precise visual feature information is less influential on letter identification when top-down orthographic information is available to compensate. Researchers have investigated this question by distorting the visual appearance of letters and measuring readers' abilities to recognise them in different letter string contexts. Letter distortion is more disruptive in single letters (Fiset et al., 2008) and pseudowords (Rosa et al., 2016) compared to real words. Therefore, existing research indicates that readers use orthographic knowledge to resolve inconsistencies in visual feature information, whether it is distortion from visual noise (Fiset et al., 2008; Rosa et al., 2016) or substitution of a visually similar letter appearing in a word-like string (e.g. *dentjst*, Marcet & Perea, 2017). However, these scenarios typically involve readers encountering an invalid string and measuring how quickly they recognise the closest word neighbour. Less is known about whether orthographic

DOI: 10.1177/17470218221106155

context reduces ambiguity from visual feature similarity if both letters result in an equally valid string. Readers often encounter this situation, as they must distinguish between word neighbours with similar looking letters (e.g. *gate-gale*). Based on previous findings, we would expect visually similar neighbours (*gate-gale)* to be harder to distinguish than visually dissimilar neighbours (*gate-game*). But how does letter confusability change across orthographic contexts? To our knowledge, researchers are yet to investigate whether orthographic context mediates readers' ability to discriminate between visually similar letters if they both result in a string with an equivalent word or non-word status.

Higher-level orthographic knowledge and low-level visual feature analysis both play a key role in letter identification, but less is known about how they interact. The visual forms of letters are highly variable; therefore, readers may use orthographic context to compensate for inconsistencies in visual feature information. Orthographic distributional knowledge provides information on how individual characters relate to each other, as readers can learn from the contexts in which letters co-occur (Schubert et al., 2020). This knowledge can reinforce mappings between variable letter shapes and identities, provide cues on the expected visual form (such as case and font), and assist in refining potential letter candidates while visual feature information is still being accumulated. These context cues not only assist readers in overcoming within-letter visual variability, but also reduce the likelihood of confusing visually similar letters. Therefore, cues from orthographic context may play a role in constraining letter candidates to manage the balance of flexibility and precision required during letter identification. If so, letter confusability from visual similarity may be reduced when wider orthographic information is available.

The focus of this work was to examine how higher-level orthographic knowledge and low-level visual feature analysis work in tandem during letter identification. We conducted a Reicher-Wheeler task to compare readers' ability to discriminate between letters with high and

DOI: 10.1177/17470218221106155

low visual feature similarity across words, pseudowords and consonant strings. We predicted that readers would be less accurate at discriminating between two letters with high visual feature overlap (*m-n*) relative to two letters with low visual overlap (*m-t*). We also predicted that letter identification would be more accurate in words relative to pseudowords, and pseudowords relative to unpronounceable consonant strings, in line with word (Reicher, 1969; Wheeler, 1970) and pseudoword superiority effects (Baron & Thurston, 1973; Carr et al., 1978). Finally, we predicted that letter confusability from visual similarity would be reduced when letter-strings aligned with orthographic and orthotactic knowledge, as we proposed that readers would use their knowledge of words and legal letter combinations to narrow down plausible letter candidates. Therefore, we predicted an interaction where accuracy differences between letters with high and low visual feature similarity would be smaller in words compared to pseudowords, and in pseudowords compared to consonant strings.

Method

Data availability

This study was pre-registered on the Open Science Framework. Pre-registration, stimuli, instructions, trial-level data, and analysis scripts are openly available (https://osf.io/p4q9u/).

Participants

Seventy-two monolingual English speakers completed the experiment at Royal Holloway University of London, in exchange for £5. All participants were aged 18-35, with normal or corrected-to-normal vision, and no previous history of reading difficulty. The sample size was determined alongside the number of items (24 items per condition) in order to meet

DOI: 10.1177/17470218221106155

the suggested criterion of 1600 observations per condition for analyses using linear mixed-effects models ($24 \times 72 = 1728$ observations per condition, Brysbaert & Stevens, 2018). All participants provided informed consent prior to taking part.

Stimuli

Target stimuli consisted of 48 words, 48 pronounceable pseudowords and 48 unpronounceable consonant strings. These three target stimuli conditions comprised the independent variable of orthographic context. Each target stimulus was assigned a target letter that was present within the stimulus, and two possible foil letters that were not present in the stimulus. Foil letters had either high visual feature overlap or low visual feature overlap with the target letter. This manipulation formed our second independent variable: visual feature similarity (high versus low). The critical target and foil letters included in visual similarity comparisons were the same across each orthographic context condition. Substitution of the target letter with either of the foil letters always resulted in a string with the same orthographic context status as the target (e.g. word: *snow/show/stow,* pseudoword: *snum/shum/stum,* consonant string: *znsq/zhsq/ztsq*). All letter strings were four to six letters long, and words and pseudowords had a single-syllable pronunciation. Word targets (*snow*) and words with the substituted foil letter (*show/stow*) were controlled for frequency using the CELEX database (Baayen et al., 1995). Stimuli for a preliminary staircase-thresholding task consisted of an additional 20 words, 20 pseudowords and 20 consonant strings, with the same control measures as those used for the main task. None of the stimuli presented in the thresholding task were present in the main task.

Visual feature similarity was quantified using seven-point letter similarity ratings from over 700 people (Simpson et al., 2013). Target letters had a mean similarity rating of 4.19 with

DOI: 10.1177/17470218221106155

foil letters in the high overlap condition compared to 1.22 with foil letters in the low overlap condition, $t(47)=24.8$, $p<.001$. This difference between high- and low-overlap conditions was confirmed with a second, objective measure of visual similarity derived from the Hierarchical Model and X (HMAX, Mutch & Lowe, 2008), a biologically motivated computational model that mimics properties of the human ventral visual system through a series of simple (S1, S2) and complex (C1, C2) layers. We used HMAX S1 layer computations to calculate letter similarities, as this layer was modelled upon the earliest instance of feature detection. HMAX calculations revealed that target letters had a mean similarity rating of 0.59 with foil letters in the high overlap condition compared to 0.50 with foil letters in the low overlap condition, $t(47)=6.25$, $p<.001$. HMAX and reader ratings were positively correlated, $r(323)=.49$, $p<.001$.

Procedure

Participants completed a Reicher-Wheeler task consisting of 144 trials, administered using DMDX (Forster & Forster, 2003). Within each trial, participants viewed a 500 ms fixation cross, followed by a forward mask for 33 ms. A target letter-string (either a word, pseudoword or consonant string) then appeared for a predetermined duration (see below), before a hash symbol (#) backward-masked each letter of the target for 100 ms. During this time, a probe bar (|) appeared above and below one of the hash symbols, which indicated that the participant should identify the letter in the specified position. After 100 ms, a target letter and a foil letter replaced the probe bars above and below one of the hash symbols. The foil letter had either high visual feature similarity or low visual feature similarity to the target letter. Participants then had 5000 ms to make a button-press response to indicate which of the two letters was present within the string. Targets were counter-balanced to ensure that participants

DOI: 10.1177/17470218221106155

received an equal number of foil letters across high and low visual feature similarity conditions, and to ensure that participants saw each target letter-string once.

Target exposure duration was determined for each participant based on performance in an initial staircase-thresholding task (adapted from Davis, 2001), which used the same trial procedure and mask durations as the main task. In the thresholding task, exposure duration began at 33 ms, and adjusted after each response. If the participant correctly identified the target letter, exposure duration was reduced by 17 ms. If the participant incorrectly identified the foil letter, exposure duration increased by 17 ms. Exposure duration was held constant after twelve changes in direction, and this value set target exposure duration for each participant in the main task. Exposure during the main experiment was 33 ms for 36 participants, 50 ms for 22 participants, 67 ms for 13 participants and 83 ms for one participant. The exposure durations were similar to previous Reicher-Wheeler studies with skilled adult readers (Chase & Tallal, 1990; Coch & Mitra, 2010; Grainger et al., 2003; Kezilas et al., 2016; Lété & Ducrot, 2008).

Results

Mean accuracy results are visualised in Figure 1. Accuracy data were analysed using logistic generalized linear mixed-effects models with the *lme4* package (Version 1.1-12; Bates et al., 2015) in *R* (Version 4.0.4; R Core Team 2016). The maximal model was defined as: glmer(Accuracy ~ Exposure Duration + (Visual Feature Similarity * Orthographic Context) + (1|Participant) + (1|Item), family=binomial). Continuous predictors (exposure duration) were centred around the mean. Categorical factor predictors (visual feature similarity and orthographic context) were dummy coded, which resulted in each level of the factor being compared to a specific level acting as a reference. For the fixed effect of visual similarity, accuracy in the high visual similarity condition was compared to accuracy in the low visual

DOI: 10.1177/17470218221106155

similarity condition as the reference. For orthographic context, accuracy performance for words and pseudowords was compared to accuracy performance for consonant strings as the reference. Therefore, the intercept of the model referred to accuracy performance within the two reference conditions (discriminating between low visual similarity letters in consonant strings).

We investigated whether each component improved the model fit using pairwise likelihood ratio tests (LRTs), in which random effects, main effects, and the interaction term were systematically added in turn (Matuschek et al., 2017). The model fit was improved by random effects of participant (LRT: $\chi^2(1)=269.76$, $p<.001$) and item (LRT: $\chi^2(1)=160.59$, $p<.001$). We then added the fixed effect of exposure duration, which referred to the duration each letter string was presented for based on participant performance in the preliminary thresholding task. The fixed effect of duration exposure continued to improve the fit of the model (LRT: $\chi^2(1)=16.75$, $p<.001$). Next, we included our fixed effects of interest. The fit of the model significantly improved after including the fixed effects of orthographic context (LRT: $\chi^2(2)=494.61$, $p<.001$) and visual feature similarity (LRT: $\chi^2(1)=32.24$, $p<.001$). However, including the interaction term did not significantly improve the model fit (LRT: $\chi^2(2)=0.36$, $p=0.838$). This indicated that there was no significant interaction between visual feature similarity and orthographic context. We opted to preserve the interaction term despite it not improving the fit, as this enabled us to test our pre-established confirmatory hypothesis that orthographic context mediates effects of visual feature similarity (see Roettger, 2019). After establishing the model fit, we ran the model and iteratively redefined the dummy-coded reference level of orthographic context to systematically compare all levels to each other. Fixed and random effects results are reported in Table 1. Beta (β) and odds ratios (*OR*) are used to report effect sizes. β is the logit transformed fixed effect coefficient, which refers to the estimated difference between conditions having controlled for random effects. *OR* (derived

DOI: 10.1177/17470218221106155

from β) measures the difference in odds of being correct (versus incorrect) in one level of a fixed effect compared to another.

--- Insert Figure 1 about here ---

--- Insert Table 1 about here ---

There was a significant main effect of visual feature similarity on letter identification accuracy. The odds ratios suggest that participants were 0.78 times as likely (or 22% less likely) to select the correct letter when the foil letter had high visual overlap with the target. There was also a significant main effect of orthographic context on letter identification accuracy. Participants were 1.86 times more likely to correctly identify the target letter in words relative to pseudowords, and 3.70 times more likely to correctly identify the letter in words relative to consonant strings. Participants were also 1.98 times more likely to correctly identify the letter in pseudowords relative to consonant strings.

There was no evidence of an interaction between visual feature similarity and orthographic context. The interaction term did not significantly improve the fit of the model and the estimated odds ratios for interaction effects were close to one (*ORs*: 0.94-0.99 or between 1-6% less likely), which indicates an equivalent likelihood of high visual similarity reducing letter identification accuracy in either orthographic context. These estimated reductions are unlikely to predict a meaningful difference, as the degree of uncertainty (shown by standard error in Table 1) indicates that each of these estimates could span either side of *OR*=1 with sampling error considered. As an additional measure, we conducted exploratory Bayesian analyses to establish whether the absence of an interaction provided evidence for the null hypothesis (i.e. that effects of visual feature similarity are not modulated by orthographic context), or whether there was insufficient evidence to infer a conclusive outcome. Using the

DOI: 10.1177/17470218221106155

motivated maximum-approach (Silvey et al., 2021, based on principles of Dienes, 2014), we calculated Bayes Factors from the interaction estimates produced by our logistic mixed-effects model. Bayes Factors (BF) were calculated using a half normal distribution (HN). Bayesian results are reported with parentheses expressing the mode of the distribution (first number), and the standard deviation (second number) in convention with Silvey et al. (2021). All BFs were lower than 0.3 (Visual Similarity * Pseudoword vs. Consonant String: $BF_{HN(0,0.11)} = 0.11$, Visual Similarity * Word vs. Consonant String: $BF_{HN(0,0.12)} = 0.17$, Visual Similarity * Word vs. Pseudoword: $BF_{HN(0,0.13)} = 0.05$). This indicated moderate evidence for the null hypothesis (Lee & Wagenmakers, 2014; Silvey et al., 2021): that orthographic context does not mediate effects of visual similarity.

Our Bayesian analyses provided evidence that there was no interaction. However, there remains a small possibility that our study was underpowered to detect it (Brysbaert, 2019a), despite our relatively large sample size (N=72) and large number of observations per condition (Brysbaert & Stevens, 2018). Thus, we ran Monte Carlo power analyses on simulated datasets to estimate the interaction effect sizes that could have been reliably detected with our sample size. Power analyses and measures taken to protect against issues of post-hoc interpretation are reported in further detail on the Open Science Framework. Using the *simr* package (Version 1.0.5; Green & MacLeod, 2016) in *R* (Version 4.0.4; R Core Team 2016), we systematically increased hypothetical interaction effect sizes by $\beta = 0.05$ and ran 50 simulations for each increment, beginning at $\beta = 0.1$. For each simulation, we modelled a larger effect between words and consonant strings relative to words and pseudowords under our hypothesis that visual similarity effects would have a greater impact on letter identification when less orthographic information is available. Our sample size (*N*=72) yielded 80% power to detect an interaction with an effect size of $\beta = 0.3$ between visual similarity differences in words and pseudowords, and an effect size of $\beta = 0.4$ between visual similarity differences in words and

DOI: 10.1177/17470218221106155

consonant strings. The equivalent odds ratios demonstrate that we had the power to detect an interaction if the benefit of having two visually distinct letters was at least 1.35 times more likely to improve letter discrimination in pseudowords relative to words, and 1.49 times more likely in consonant strings relative to words. These analyses show that, if there was an undetected interaction between visual similarity and orthographic context in our data, it was smaller than the effect sizes stated above.

Discussion

Our results revealed effects of orthographic context and visual feature similarity on letter discrimination accuracy in a Reicher-Wheeler task. Performance improved as letter strings became more word-like (words > pseudowords > consonant strings), replicating the word superiority effect and the pseudoword superiority effect (Baron & Thurston, 1973; Carr et al., 1978; Reicher, 1969; Wheeler, 1970). Performance was also superior when the discrimination involved letters with low visual similarity compared to letters with high visual similarity. There was no interaction between the effects of orthographic context and visual feature similarity; visually similar letters were more confusable irrespective of how closely the target letter string aligned with a real word. Odds ratios indicated that effects of orthographic context were much larger than effects of visual feature similarity, which suggests that top-down orthographic knowledge may be relatively more important than bottom-up feature information in establishing letter identities.

We had hypothesised that there would be an interaction between visual similarity and orthographic context, such that effects of visual feature similarity would be stronger where there is less higher-level orthographic information available. However, there was no evidence to suggest that the impact of visual similarity on letter confusability varied across word, pseudowords or consonant strings. Exploratory Bayesian analyses indicated moderate evidence

DOI: 10.1177/17470218221106155

against an interaction, although increased power could provide the benefit of greater certainty. Power simulations revealed that our design was not powered to detect the interaction effect sizes revealed in the model output (Table 1); thus, these estimates of the interaction effect size need to be treated with caution. However, detecting an interaction effect of this size would require a sample of at least 5,000 participants (based on 80% power, calculations available on the OSF). It may be more constructive in future research to determine what an ecologically meaningful interaction effect size would be and calculate power accordingly.

Our findings advance current understanding of letter identification in several ways. First, we believe this work to be one of the first to demonstrate effects of visual feature similarity on letter identification within a Reicher-Wheeler paradigm. This departs from previous work investigating visual feature similarity, which has mostly been restricted to masked priming (although see also Marcet & Perea, 2018b, for parafoveal preview effects). The current work demonstrates that effects of visual feature similarity are not task-specific, and have implications for multiple levels of processing. Pre-existing evidence from masked-priming demonstrated that visual feature similarity has a discernible influence on low-level perceptual processing (i.e. visual similarity between the prime and the target) and processes that rely upon broad lexical knowledge (i.e. visual similarity between the prime and known word strings), whereas the current findings show that visual feature similarity also impacts processing when readers are required to specifically discriminate between letter candidates.

Second, this study has taken a new approach to investigating visual similarity effects, by investigating the impact of shared featural information across letters that result in equally valid letter strings (i.e. both letters result in a real word, for example). In masked-priming paradigms, researchers have typically compared visual overlap between pseudoword and word neighbours, whereby the pseudoword is the prime and the word neighbour is the target (e.g. *dentjst-dentist* vs. *dentgst-dentist*; Marcet & Perea, 2017). This has yielded powerful initial evidence that

DOI: 10.1177/17470218221106155

shared featural information can be beneficial for visual word recognition, as readers are faster at overcoming discrepancies between a letter string and the closest known word form if the two are visually similar. When relating to real-life reading experience, this advantage is akin to recovering from a typing or spelling mistake. In contrast, the current work investigates an alternative problem, as it is one of the first to investigate the influence of visual similarity when discriminating between letters that result in equally plausible word forms. In this scenario, readers are not assessing overlap between the visual input and an expected letter form, but instead distinguishing between competing letter identities. This is an alternative but also common challenge during reading, as readers routinely discriminate between word neighbours with similar looking letters (e.g. *gate*-*gale*). This critical difference provided a new insight: visual feature similarity may benefit visual word recognition when one letter is more likely to occur than the other, but it can also be disadvantageous if both letters are equally plausible, as it hinders readers ability to discriminate between them. This suggests that readers refine potential letter candidates whilst visual feature information is still being accumulated. Visual feature similarity impedes visual word recognition when both letters are equally viable, as neither competitor has been disregarded as an unsuitable candidate.

This interpretation is further supported by an additional novel conclusion from this work, which relates to how cues are weighted during letter identification. The influence of orthographic context was much larger than the influence of visual feature similarity, which suggests that top-down orthographic knowledge may be prioritised over bottom-up feature information during letter identification. This differential weighting may again occur because orthographic knowledge plays a critical role in filtering letter candidates, enabling readers to maintain the balance of flexibility and precision required for letter identification. Readers must incorporate a certain degree of flexibility when mapping low-level visual features to letter identities, as the visual appearance of letters can be highly variable. However, allowing greater

DOI: 10.1177/17470218221106155

flexibility also increases the risk of letter confusability. We propose that orthographic knowledge mitigates this risk while visual feature information is still being accumulated, by disregarding unlikely letter candidates and prioritising those that would result in a real word or an orthotactically legal letter string.

This finding has important theoretical implications, as understanding the weighting attributed to word-level (or 'string-level') cues relative to visual feature cues is essential for understanding the integration of visual and linguistic information. This proposed 'mid-level vision stage' of orthographic processing (Grainger, 2018) is often under-specified in cognitive models of reading, as feature-level processes are either assumed *a priori* or minimally outlined (Marcet & Perea, 2017). There has been greater focus in neuro-biological models of reading, which incorporate mechanisms for visual object identification to interact with linguistic processing in order to facilitate visual word recognition. For example, the local combination detector model (Dehaene et al., 2005) proposes that readers hierarchically encode increasingly large fragments of orthographic information that advance in linguistic complexity (features, letter shapes, abstract letter identities, bigrams, substrings) in the visual ventral stream, with increasing sensitivity to linguistic probabilities within the writing system. The model outlines how feature-level information may be incorporated based on principles of the primate visual system; readers amalgamate oriented bars and local contours to detect letter shapes, which are then used to inform abstract letter representations invariant of font or case. Thus, the local combination detector model is able to explain why visually similar letters are more confusable. However, this model assumes a one-way feed-forward approach, which restricts its ability to explain how orthographic knowledge reduces letter confusability. Without incorporating feedback, it is difficult to align this account with our finding that orthographic status has a much larger influence on letter confusability than low-level visual similarity, particularly as this was observed in a Reicher-Wheeler task with limited exposure to the visual input.

DOI: 10.1177/17470218221106155

Alternatively, the influence of orthographic knowledge on refining letter candidates can be characterised by principles of cascaded processing, whereby later stages of word processing are implemented before earlier stages are completed (McClelland, 1979). Cascaded processing can explain why effects of visual feature similarity are outweighed by cues from lexical information when available, as word-level feedback plays a greater role in activating letter representations compared to bottom-up activation from feature-level information alone. Recent neural evidence has indicated how cascaded processing may be incorporated into existing neuro-biological models, following detection of feedforward and feedback activity within the ventral stream (Woolnough et al., 2020). Woolnough et al. (2020) found that posterior regions were the earliest to show increased activation in response to orthographic stimuli, however, these regions also demonstrated sensitivity to lexical status later than anterior regions. Differences in early and late selectivity could reflect cascaded processing, as word-likeness recognised in anterior regions may propagate backwards and interact with letter-level processing in posterior regions. Thus, there is potential to inform a cascaded model which could incorporate both the analysis of visual information and feedback from linguistic knowledge.

The greater weighting attributed to higher-level orthographic information could otherwise potentially be explained by Bayesian models of reading, which propose that visual word recognition is achieved by readers combining tentative evidence with knowledge of prior probability (Norris, 2006; Norris et al., 2010; Norris & Kinoshita, 2012). Under this interpretation, bottom-up analysis of lower-level orthographic features constitutes the tentative evidence and integration of top-down orthographic knowledge shapes the priors of the expected visual word representation. The greater influences of higher- level orthographic cues (i.e. word status) relative to lower-level visual cues (i.e. feature information) may be due to readers having stronger priors for letter combinations associated with known word representations or phonotactically legal letter combinations, which requires less detailed analysis of the visual

DOI: 10.1177/17470218221106155

evidence. It is less clear how these principles would be ingrained in visual processing, although there is again potential to consider the compatibility of these principles with existing neuro-biological or visual models of reading. For example, previous neuro-imaging work has documented how top-down predictions influence the sensory processing of speech (Davis & Sohoglu, 2020; Sohoglu et al., 2012). Future work could investigate similar principles for reading within the visual domain.

In conclusion, the current study demonstrated that letter identification is supported through a balance of information from visual features and higher-level orthographic knowledge. Our results showed that visually similar letters are more confusable than dissimilar letters, indicating that readers encode letter identities with initial uncertainty, based on feature information. Word and pseudoword superiority effects demonstrated that readers also use orthographic knowledge of known words and legal letter combinations to resolve early uncertainty around letter identity. In the absence of an interaction, there is no evidence to suggest that orthographic context mediates effects of visual similarity specifically. Instead, our findings indicate that orthographic knowledge and visual feature similarity have an additive effect on letter identification. This work provides a novel insight that higher-level orthographic information plays a greater role than lower-level visual feature information in letter identification. We suggest that this is a result of readers using orthographic knowledge to refine potential letter candidates efficiently with while visual feature information is still being accumulated. More broadly, this work advances understanding of the integration of visual and linguistic information and highlights the need for greater cross-examination between models of reading and visual processing.

DOI: 10.1177/17470218221106155

Data Accessibility Statement

The data and materials from the present experiment are publicly available at the Open Science Framework website: https://osf.io/p4q9u/

DOI: 10.1177/17470218221106155

References

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1995). The CELEX Lexical Database. Release 2 [CD-ROM]. Linguistic Data Consortium, University of Pennsylvania, Philadelphia.

Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology*, *4*(2), 207-228. https://doi.org/10.1016/0010-0285(73)90012-1.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R Package Version 1.1–7. 2014.

Bowers, J. S., Vigliocco, G., & Haan, R. (1998). Orthographic, phonological, and articulatory contributions to masked letter and word priming. *Journal of Experimental Psychology: Human Perception and Performance,* *24*(6), 1705–1719. https://doi.org/10.1037/0096-1523.24.6.1705.

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1). http://doi.org/10.5334/joc.72.

Brysbaert, M ., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: a tutorial. *Journal of Cognition. 1*(1). http://doi.org/10.5334/joc.10.

Carr, T. H., Davidson, B. J., & Hawkins, H. L. (1978). Perceptual flexibility in word recognition: Strategies affect orthographic computation but not lexical access. *Journal of Experimental Psychology: Human Perception and Performance,* *4*(4), 674–690. https://doi.org/10.1037/0096-1523.4.4.674.

Chase, C. H., & Tallal, P. (1990). A developmental, interactive activation model of the word superiority effect. *Journal of Experimental Child Psychology, 49*(3), 448-487. https://doi.org/10.1016/0022-0965(90)90069-K.

DOI: 10.1177/17470218221106155

Coch, D., & Mitra, P. (2010). Word and pseudoword superiority effects reflected in the ERP waveform. *Brain Research*, *1329*, 159-174. https://doi.org/10.1016/j.brainres.2010.02.084.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. https://doi.org/10.1037/0033-2909.112.1.155.

Coltheart, M., Rastle, K., Perry, C., Langdon, R ., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204-256. https://doi.org/10.1037/0033-295x.108.1.204.

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, M.H. (2001). *Up-Down Staircase Psychophysical Method.* Retrieved from http://www.u.arizona.edu/~kforster/dmdx/up-down_staircase.htm.

Davis, M. H., & Sohoglu, E. (2020). Three functions of prediction error for Bayesian inference in speech perception. In M. Gazzaniga, R. Mangun, & D. Poeppel (Eds.), *The Cognitive Neurosciences, 6th Edition,* 177-189. MIT Press.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., ... & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, *330*(6009), 1359-1364. https://doi.org/10.1126/science.1194140.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5, 1–17. https://doi.org/10.3389/fpsyg.2014.00781.

Dziak, J. J., Dierker, L. C., & Abar, B. (2020). The interpretation of statistical power after the data have been gathered. *Current Psychology*, *39*(3), 870-877. https://doi.org/10.1007/s12144-018-0018-1.

DOI: 10.1177/17470218221106155

Faul, F., Erdfelder, E., Lang, A. G ., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146.

Fiset, D., Blais, C., Ethier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for identification of uppercase and lowercase letters. *Psychological Science*, *19*(11), 1161-1168. https://doi.org/10.1111/j.1467-9280.2008.02218.x.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behaviour Research Methods*, *35(1),* 116-124. https://doi.org/10.3758/BF03195503.

Goldstein, A. (1964). *Biostatistics: An introductory text*. New York: MacMillan.

Grainger, J., Bouttevin, S., Truc, C., Bastien, M., & Ziegler, J. (2003). Word superiority, pseudoword superiority, and learning to read: A comparison of dyslexic and normal readers. *Brain and Language*, *87*(3), 432-440. https://doi.org/10.1016/S0093-934X(03)00145-7.

Grainger, J., Dufau, S., & Ziegler, J. C. (2016). A vision of reading. *Trends in Cognitive Sciences*, *20*(3), 171-179. https://doi.org/10.1016/j.tics.2015.12.008.

Grainger, J., & Jacobs, A. M. (1994). A dual read-out model of word context effects in letter perception: Further investigations of the word superiority effect. *Journal of Experimental Psychology: Human Perception and Performance,* *20*(6), 1158–1176. https://doi.org/10.1037/0096-1523.20.6.1158.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*(3), 518–565. https://doi.org/10.1037/0033-295X.103.3.518.

DOI: 10.1177/17470218221106155

Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493-498. https://doi.org/10.1111/2041-210X.12504.

Hannagan, T., Ktori, M., Chanceaux, M., & Grainger, J. (2012). Deciphering CAPTCHAs: What a Turing test reveals about human cognition. *PLOS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0032121.

Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power: the pervasive fallacy of power calculations for data analysis. *The American Statistician*, *55*(1), 19-24. https://doi.org/10.1198/000313001300339897.

Kezilas, Y., Kohnen, S., McKague, M., Robidoux, S., & Castles, A. (2016). Word and pseudoword superiority effects on letter position processing in developing and skilled readers. *Journal of Experimental Psychology: Human Perception and Performance, 42*(12), 1989-2002. https://doi.org/10.1037/xhp0000273.

Kinoshita, S ., & Kaplan, L. (2008). Priming of abstract letter identities in the letter match task. *The Quarterly Journal of Experimental Psychology*, *61*(12), 1873-1885. https://doi.org/10.1080/17470210701781114.

Kinoshita, S., Robidoux, S., Mills, L ., & Norris, D. (2013). Visual similarity effects on masked priming. *Memory & Cognition*, *42*(5), 821-833. https://dx.doi.org/10.3758%2Fs13421-013-0388-4.

Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2020). Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of Gerontology: Series B*, *75*(1), 45-57. https://doi.org/10.1093/geronb/gby065.

DOI: 10.1177/17470218221106155

Lee, M. D., & Wagenmakers, E.J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Lété, B., & Ducrot, S. (2008). Visuo-attentional processing by dyslexic readers in the Reicher–Wheeler task. *Current Psychology Letters, 24*, http://www.cpl.revues.org/document3523.html.

Lien, M. C., Allen, P., & Martin, N. (2014). Processing visual words with numbers: Electrophysiological evidence for semantic activation. *Psychonomic Bulletin & Review*, *21*(4), 1056-1066. https://doi.org/10.3758/s13423-014-0581-x.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 476-490. http://doi.org/10.3758/BF03210951.

Marcet, A., & Perea, M. (2017). Is nevtral NEUTRAL? Visual similarity effects in the early phases of written-word recognition. *Psychonomic Bulletin & Review*, *24*(4), 1180-1185. https://doi.org/10.3758/s13423-016-1180-9.

Marcet, A., & Perea, M. (2018a). Can I order a burger at rnacdonalds. com? Visual similarity effects of multi-letter combinations at the early stages of word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(5), 699-706. https://doi.org/10.1037/xlm0000477.

Marcet, A., & Perea, M. (2018b). Visual letter similarity effects during sentence reading: Evidence from the boundary technique. *Acta Psychologica*, *190*, 142-149. https://doi.org/10.1016/j.actpsy.2018.08.007.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305-315. https://doi.org/10.1016/j.jml.2017.01.001.

DOI: 10.1177/17470218221106155

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375-407. https://doi.org/10.1037/0033-295X.88.5.375.

Mutch, J., & Lowe, D. G. (2008). Object class recognition and localization using sparse features with limited receptive fields. *International Journal of Computer Vision*, *80*(1), 45-57. https://doi.org/10.1007/s11263-007-0118-0.

Norris, D. (2006). The Bayesian reader: explaining word recognition as an optimal Bayesian decision process. *Psychological Review, 113*(2), 327-57. https://doi.org/10.1037/0033-295x.113.2.327.

Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review*, *119*(3), 517-545. https://psycnet.apa.org/doi/10.1037/a0028450.

Norris, D., Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language*, *62*(3), 254-271. https://doi.org/10.1016/j.jml.2009.11.002.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*(28), 4646-4674. https://doi.org/10.1016/j.visres.2006.04.023.

Pelli , D., Majaj, N. J., Raizman, N., Christian, C.J., Kim, E ., & Palomares, M.C. (2009) Grouping in object recognition: The role of a Gestalt law in letter identification, *Cognitive Neuropsychology*, *26*(1), 36-49. https://doi.org/10.1080/13546800802550134.

Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). R34d1ng w0rd5 w1th numb3r5. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 237-241. https://doi.org/10.1037/0096-1523.34.1.237.

DOI: 10.1177/17470218221106155

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modelling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology*, *61*(2), 106-151. https://doi.org/10.1016/j.cogpsych.2010.04.001.

Petit, J. P., & Grainger, J. (2002). Masked partial priming of letter perception. *Visual Cognition*, *9*(3), 337-353. https://doi.org/10.1080/13506280042000207.

R Core Team (2016). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from: http://www.R-project.org/.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*(2), 275-280. https://doi.org/10.1037/h0027768.

Roettger, T., (2019) Researcher degrees of freedom in phonetic research, *Laboratory Phonology* 10(1), 1, https://doi.org/10.5334/labphon.147.

Rosa, E., Perea, M., & Enneson, P. (2016). The role of letter features in visual-word recognition: Evidence from a delayed segment technique. *Acta Psychologica*, *169*, 133-142. https://doi.org/10.1016/j.actpsy.2016.05.016.

Rothlein, D., & Rapp, B. (2014). The similarity structure of distributed neural responses reveals the multiple representations of letters. *NeuroImage*, *89*, 331-344. https://doi.org/10.1016/j.neuroimage.2013.11.054.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, *89*(1), 60–94. https://doi.org/10.1037/0033-295X.89.1.60.

DOI: 10.1177/17470218221106155

Rumelhart, D. E., & Siple, P. (1974). The process of recognizing tachistoscopically presented words. Psychological Review. *81*(2), 99–118. https://psycnet.apa.org/doi/10.1037/h0036117.

Schubert, T. M., Cohen, T., & Fischer-Baum, S. (2020). Reading the written language environment: Learning orthographic structure from statistical regularities. *Journal of Memory and Language*, *114*, 104148. https://doi.org/10.1016/j.jml.2020.104148

Silvey, C., Dienes, Z., & Wonnacott, E. (2021). Bayes factors for mixed-effects models. *PsyArXiv.* https://doi.org/10.31234/osf.io/m4hju.

Simpson, I. C., Mousikou, P., Montoya, J. M., & Defior, S. (2013). A letter visual-similarity matrix for Latin-based alphabets. *Behavior Research Methods*, *45*(2), 431-439. https://doi.org/10.3758/s13428-012-0271-4.

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience*, *32*(25), 8443-8453. https://doi.org/10.1523/JNEUROSCI.5069-11.2012.

Wheeler, D. (1970). Processes in word recognition. *Cognitive Psychology*, *1*(1), 59-85. https://doi.org/10.1016/0010-0285(70)90005-8.

Wiley, R. W., Wilson, C., & Rapp, B. (2016). The effects of alphabet and expertise on letter perception. *Journal of Experimental Psychology: Human Perception and Performance, 42*(8), 1186–1203. https://doi.org/10.1037/xhp0000213.

DOI: 10.1177/17470218221106155

Figure Captions

Figure 1. Mean accuracy rates for letter identification in the Reicher-Wheeler task. Crossbars display mean accuracy rates across participants and tiles display one standard error from the mean, calculated for within-subject designs (Loftus & Masson, 1994). Data points display accuracy rates for individual participants and violins demonstrate the distribution of the data. The dashed horizontal line displays chance performance.

DOI: 10.1177/17470218221106155

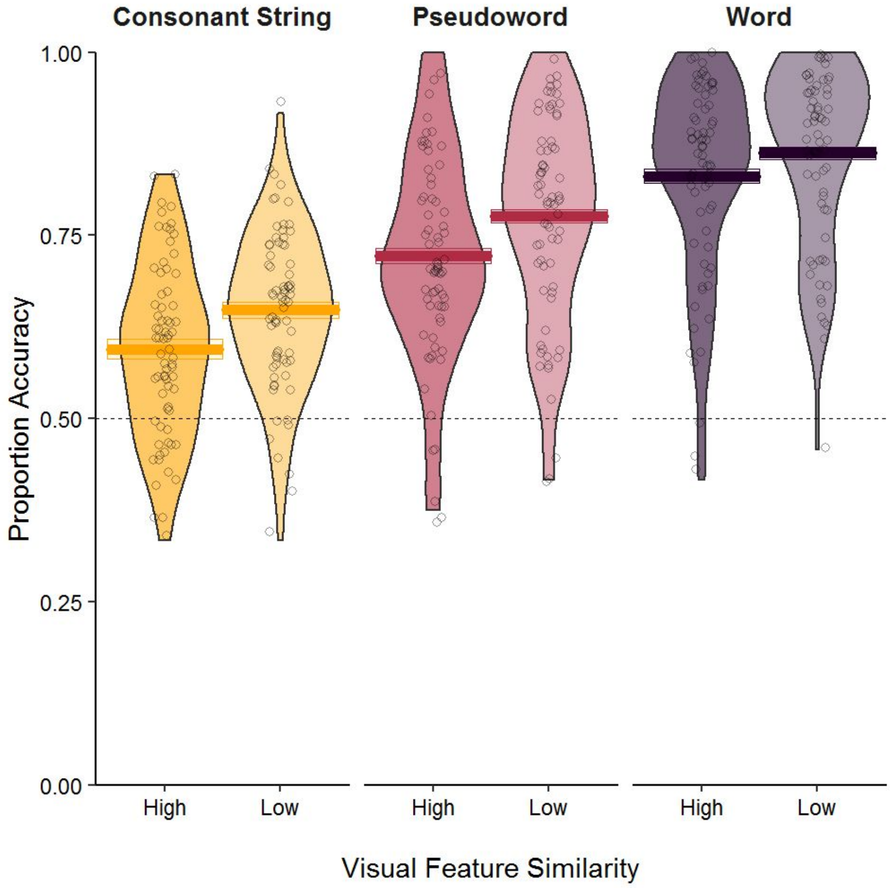

Figure 1. Mean accuracy rates for letter identification in the Reicher-Wheeler task. Crossbars display mean accuracy rates across participants and tiles display one standard error from the mean, calculated for within-subject designs (Loftus & Masson, 1994). Data points display accuracy rates for individual participants and violins demonstrate the distribution of the data. The dashed horizontal line displays chance performance.

DOI: 10.1177/17470218221106155

Table 1. Logistic generalized linear mixed effects model output for analysis of exposure duration, visual feature similarity and orthographic context on letter identification. Beta and odds ratio effect sizes are reported, with standard error in parentheses.

Fixed Effects	**β *(SE)***	***OR (SE)***	***z***	***p***
Exposure Duration	0.02 (0.00)	1.02 (0.00)	4.31	<0.001
Visual Feature Similarity: High vs. Low	-0.25 (0.07)	0.78 (0.06)	-3.37	<0.001
Orthographic Context: Pseudoword vs. Consonant String	0.68 (0.08)	1.98 (0.16)	8.62	<0.001
Orthographic Context: Word vs. Consonant String	1.31 (0.09)	3.70 (0.33)	14.74	<0.001
Orthographic Context: Word vs. Pseudoword	0.62 (0.09)	1.86 (0.17)	6.70	<0.001
Visual Similarity * Orthographic Context: Pseudoword vs. Consonant String	-0.06 (0.11)	0.94 (0.10)	-0.57	0.569
Visual Similarity * Orthographic Context: Word vs. Consonant String	-0.01 (0.12)	0.99 (0.12)	-0.07	0.946
Visual Similarity * Orthographic Context: Pseudoword vs. Word	-0.05 (0.13)	0.95 (0.12)	-0.43	0.669
Random Effects				
σ2	3.29		**τ00**	0.19 $_{\text{SubID}}$
ICC	0.09			0.15 $_{\text{Item}}$
Observations	10368		**N**	72 $_{\text{SubID}}$
Marginal R2 / Conditional R2	0.091 / 0.176			48 $_{\text{Item}}$

DOI: 10.1177/17470218221106155