Journal of Experimental Psychology:
Learning, Memory, and Cognition
2018, Vol. 44, No. 5, 699–706

0278-7393/18/$12.00 http://dx.doi.org/10.1037/xlm0000477

Can I Order a Burger at rnacdonalds.com? Visual Similarity Effects of Multi-Letter Combinations at the Early Stages of Word Recognition

Ana Marcet and Manuel Perea
Universitat de València

Previous research has shown that early in the word recognition process, there is some degree of uncertainty concerning letter identity and letter position. Here, we examined whether this uncertainty also extends to the mapping of letter features onto letters, as predicted by the Bayesian Reader (Norris & Kinoshita, 2012). Indeed, anecdotal evidence suggests that nonwords containing multi-letter homoglyphs (e.g., rn→m), such as *documnent*, can be confusable with their base word. We conducted 2 masked priming lexical decision experiments in which the words/nonwords contained a middle letter that was visually similar to a multi-letter homoglyph (e.g., *docurnent* [rn–m], *presiclent* [cl–d]). Three types of primes were employed: identity, multi-letter homoglyph, and orthographic control. We used 2 commonly used fonts: Tahoma in Experiment 1 and Calibri in Experiment 2. Results in both experiments showed faster word identification times in the homoglyph condition than in the control condition (e.g., *docurnento–DOCUMENTO* faster than *docusnento–DOCUMENTO*). Furthermore, the homoglyph condition produced nearly the same latencies as the identity condition. These findings have important implications not only at a theoretical level (models of printed word recognition) but also at an applied level (Internet administrators/users).

Keywords: lexical access, masked priming, visual similarity, models of word recognition

In Roman script, as in other alphabetic scripts, letters are the building blocks of words. When a printed word is presented, the initial stages of processing are devoted to identify the visual features (e.g., oriented curves/lines, intersections, terminations, among others) of each of the visual objects (i.e., letters) that comprise the word (e.g., the letters s-a-l-t in the word *salt*; see Figure 1). The activation from these visual feature detectors is sent to abstract letter detectors (i.e., *salt* and *SALT* would activate the same letter detectors), which, in turn, send activation to whole-word units; note that this process may involve both feedforward and feedback connections (Carreiras, Armstrong, Perea, & Frost, 2014). The "magic moment" in word recognition occurs when the level of activation of a given whole-word unit exceeds some threshold (see Balota, Yap, & Cortese, 2006; McClelland, Mirman, Bolger, & Khaitan, 2014).

Thus, as Norris and Kinoshita (2012) pointed out, during printed word recognition "the visual system must identify how many visual objects are present, their configuration (order), and their identity" (p. 521). Indeed, the cognitive processes that underlie the encoding of letter identification and letter position have received considerable attention in the literature on printed word recognition and reading (see Grainger, Dufau, & Ziegler, 2016, for review). Most researchers assume that there is some degree of uncertainty regarding letter identity and letter position in the early moments of word processing (e.g., see Davis, 2010; Gomez, Ratcliff, & Perea, 2008; Grainger & van Heuven, 2003; Norris, 2006). For instance, using Forster and Davis's (1984) masked priming technique (i.e., a technique devised to reflect the early stages of word processing), Marcet and Perea (2017) found that word identification times on a target stimulus (e.g., *DENTIST*) were shorter when preceded by a visually similar one-letter replacement nonword prime (e.g., *dentjst*; note that "j" and "i" are visually similar) than when preceded by a visually dissimilar one-letter replacement nonword prime (e.g., *dentgst*; see Kinoshita, Robidoux, Mills, & Norris, 2014; Perea, Duñabeitia, & Carreiras, 2008, for similar evidence using letter-like numbers [e.g., 4 = A in *M4TERI4L*] and symbols [e.g., Δ = A in *MΔTERIΔL*]). These findings suggest that, at the earliest stages of word processing, the letter detector corresponding to the letter "A" can be activated not only by the sensory representation of the letter "A" but also by visually similar characters such as "H," "4," or "Δ." Likewise, word recognition times on a target stimulus (e.g., *JUDGE*) are shorter when preceded by a transposed-letter nonword prime (e.g., *jugde*) than when preceded by a replacement-letter nonword prime (e.g., *jupte*; Perea & Lupker, 2003, 2004; see Johnson, Perea, & Rayner, 2007, for evidence during sentence reading using parafoveal previews). The robustness of masked transposed-letter priming effects (e.g., *jugde* → *judge* or *cholocate* → *chocolate*) is a demonstration that, early in processing, there is some degree of ambiguity regarding letter position.

This article was published Online First November 2, 2017.

Ana Marcet and Manuel Perea, Departamento de Metodología and ERI Lectura, Universitat de València.

The research reported in this article has been partially supported by Grants PSI2014-53444-P and BES-2015-07414 from the Spanish Ministry of Economy and Competitiveness. We thank Jeff Bowers, Steve Lupker, and Jingjing Zhao for constructive comments on a previous version of this article.

Correspondence concerning this article should be addressed to Manuel Perea, Departamento de Metodología, Avenida Blasco Ibáñez, 21, 46010-Valencia, Spain. E-mail: mperea@uv.es

Figure 1. Representation of the word *salt* in Calibri, indicating the width of each letter slot. See the online article for the color version of this figure.

Critically, much less attention has been dedicated to the perceptual front end of word processing: how the visual objects that compose the words are mapped onto letters. As Finkbeiner and Coltheart (2009) indicated, "determining how readers extract letter identities from the highly variable featural information in the input is fundamental to attempts to understand the reading process" (p. 2). However, none of the current computational models of printed word recognition and reading make any specific claims on the binding from letter feature to letters. This is so because (a) the focus of these models is on already highly complex processes (i.e., the underpinnings of the letter and word levels), and (b) the "basic results" are assumed to be independent of the implementation of a detailed mapping between features and letters (see Davis, 2010, p. 725; McClelland & Rumelhart, 1981, p. 383). As an illustrative example, when introducing the SERIOL model of visual word recognition, Whitney (2001) stated the following: "We do not model the process of letter recognition; we take as given that a mechanism exists to bind the features of a letter together, culminating in activation of the correct letter" (pp. 226–227).

For simplicity, the interactive activation model (McClelland & Rumelhart, 1981) and its successors (dual-route cascaded model: Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; multiple read-out model: Grainger & Jacobs, 1996) assume discrete slots for letters that are fed by a feature-level channel on each letter position using the Rumelhart and Siple (1974) uppercase font (e.g., the word *HAND* would be encoded as ). The model does not make any claims on how the visual components of letters are put together, although a reasonable starting point is that the small gaps among letters act as boundaries (e.g., we may not know Georgian script, but we may easily deduce that the word საღამო [*night* in Georgian] is composed of six letters). Therefore, the orthographic coding scheme of the interactive activation model predicts an unequivocal mapping of letter features to letters (e.g., the letter features that compose "H" in *HAND* would not be merged with the letter features of A).[1] However, as indicated earlier, there is a rich literature that has repeatedly shown that the processes underlying word recognition can be better understood assuming a noisy or incomplete signal of the visual input at the early moments of processing (Adelman, 2011; Davis, 2010; Gomez et al., 2008; Grainger et al., 2016; Norris, Kinoshita, & van Casteren, 2010).

Therefore, an open question is whether there is some degree of ambiguity in the initial mapping of letter features onto letters. Importantly, a leading model of visual word recognition (Bayesian Reader model; Norris & Kinoshita, 2012) makes specific predictions in this respect. Norris and Kinoshita (2012) claimed that, in the early moments of word processing, "there will be uncertainty about the identity of the objects, their location, and even whether the objects really exist or are insertions created by spurious noise in the system" (p. 521)—this ambiguity would be progressively resolved over time (i.e., *JUGDE* may be processed initially as *JUDGE*, but at some point the reader will notice the difference; see Vergara-Martínez, Perea, Gómez, & Swaab, 2013, for an analysis of the time course of the transposed-letter effect using evoked-related potentials). Thus, the Bayesian Reader model can successfully capture that the masked nonword prime *dentjst* is more effective at activating the word *DENTIST* than the control nonword *dentgst* (i.e., *dentjst* and *dentist* only differ in a visually similar letter) and that the masked nonword prime *jugde* is more effective at activating the word *JUDGE* than the control nonword *jupte* (i.e., *jugde* and *judge* share the same letters in different order). But more importantly for the present purposes, the Bayesian Reader model predicts that early in word processing, there is also some degree of uncertainty with respect to the mapping of visual objects onto letters. As Norris and Kinoshita (2012) indicated, "In the same way that we assume that identity and order information accumulates gradually over time, we also assume that knowledge of which letters, or letter objects, are in the input also improves over time" (p. 524).

One way to test this assumption of the Bayesian Reader model is to examine whether or not visually presented words are segmented into letter units at the early stages of word processing (i.e., *house*: h-o-u-s-e). To tackle this issue, we took advantage of the fact that there are multi-letter combinations—the so-called "multiletter homoglyphs"[2]—whose shapes may resemble individual letters (e.g., rn → m, cl → d, vv → w). Although this type of confusion occurs in optical character recognition engines (e.g., the Tesseract ORG engine [Smith, 2007] has a specific module to avoid "rn" being identified as "m"), we must keep in mind that human readers can rapidly read words composed of distorted characters (e.g., Completely Automated Public Turing test to tell Computers and Humans Apart [CAPTCHAs]; see von Ahn, Maurer, McMillen, Abraham, & Blum, 2008) that pose problems to OCR engines (see Hannagan, Ktori, Chanceaux, & Grainger, 2012, for evidence of substantial masked repetition priming when using CAPTCHAs as primes). Thus, the research question could be put this way: Would the detectors of a given letter (e.g., "m") be activated early in processing when a word not containing this specific letter—but containing a multi-letter homoglyph (e.g., *documnent*)—is briefly presented? This question is not only important at a theoretical level (i.e., it would help refine the perceptual feature-letter front-end of models of printed word recognition); at an applied level, the potential confusability across letters is a matter of serious concern when accessing Internet websites (Davis & Suignard, 2012; see also Bohm, 2015). As acknowledged on the

[1] The story is (obviously) more complex for semicursive scripts (e.g., Arabic; see Yakup, Abliz, Sereno, & Perea, 2015). For instance, in the Arabic sentence "أخذ الولد دب صديقه" [*the boy took the friend bear* in Arabic], each word's segments need to be adequately segmented into letters and words. The same case applies to reading cursive handwriting, in which the lack of uniformity across letters adds to the additional segmentation processes (e.g., see Barnhart & Goldinger, 2010; Perea, Marcet, Uixera, & Vergara-Martínez, 2017).

[2] As defined in Wikipedia, "A homoglyph is one of two or more graphemes, characters, or glyphs with shapes that appear identical or very similar" (Homoglyph, n.d.).

Microsoft website when discussing security in Internet domain names, "rnicrosoft.com looks much like microsoft.com,"[3] and this may lead to spoofing attacks (see Gabrilovich & Gontmakher, 2002; Krammer, 2006). Thus, leaving aside the theoretical implications of the processing of multi-letter homoglyphs, if the suspicion that these letter combinations are processed as one letter is confirmed empirically, great care should be taken to prevent scammers from imitating false identity via multi-letter homoglyphs—as in *rnacdonals.com*— on Internet websites.

In the present experiments, we selected 240 words composed of a middle letter that resembled a multi-letter homoglyph (e.g., "rn" and "m" [*docurnent–document*]; "cl" and "d" [*presiclent–president*])—the multi-letter homoglyph "vv" was not used because the letter "w" is very infrequent in Spanish. As we were interested in examining the early stages of word recognition, we used the same procedure as in prior research on letter identity and letter position coding (i.e., masked priming lexical decision; see Marcet & Perea, 2017; Perea & Lupker, 2004). On each trial, an uppercase target stimulus (e.g., *DOCUMENTO* [the Spanish for *DOCUMENT*]) was preceded by a 50-ms lowercase nonword prime created by replacing the critical letter (e.g., "m") of the target word (*documento*) by its corresponding multi-letter homoglyph (*docurnento*) or by a lowercase nonword control prime created by replacing the letters "rn" with "sn" (*docusnento*).[4] To obtain an estimate of the degree that the multi-letter homoglyphs activate their visually similar equivalent letters, we also included a lowercase identity prime (*documento*).

The predictions are clear. If visually presented words are readily segmented into letter units at early stages of word processing, the letters "r" and "n" in *docurnent* would only activate their corresponding best-match letter units (i.e., "r" and "n"). As a result, the multi-letter homoglyph "rn" would not activate the letter "m" in *docurnent* to a greater degree than the control multi-letter combination "sn" in *docusnent*. Therefore, one would expect a similar advantage of the identity condition (*document–DOCUMENT*) over the two replacement-letter conditions (*docurnent–DOCUMENT* and *docusnent–DOCUMENT*). At a theoretical level, this outcome would support the idea that the small gaps around the visual objects mark the beginning and end of each letter, at least when using highly legible fonts; furthermore, this would pose some problems to those models that assume that early in processing, there is some degree of uncertainty at assigning the word's constituent visual objects onto letter units. At an applied level, this outcome would suggest that the source of the alleged confusability of multi-letter homoglyphs (if any) does not arise early in processing. Alternatively, if there were some degree of uncertainty at assigning the stimulus' visual information to letters units early in processing (i.e., the letter "rn" in *docurnento* would activate the abstract unit "m" to some degree)—as would predict the Bayesian Reader model (Norris & Kinoshita, 2012)— one would expect faster word identification times in the multi-letter homoglyph priming condition than in the control priming condition (i.e., *docurnent–DOCUMENT* < *docusnent–DOCUMENT*); indeed, in the extreme scenario, the multi-letter homoglyph condition could be processed as fast as the identity condition (for evidence with visually similar one-letter different primes, see Marcet & Perea, 2017; Perea et al., 2008). This latter outcome would not only support the predictions of the Bayesian Reader model (Norris & Kinoshita, 2012), but it would make it necessary to refine the links between the feature and letter levels in models of printed word recognition—note that the Bayesian Reader model is silent as to how the visual objects are bound onto letters. (We defer a discussion how visual features are mapped onto letters to the General Discussion.) Furthermore, at an applied level, this outcome should be made known to Internet administrators to avoid users from being potential victims of identity thief on malicious websites—this may also lead to the creation and use of fonts that minimize this potential confusability.

To assess the generality of the findings in an ecological setting, we employed two common fonts. In Experiment 1, we used Tahoma, which is the default font of the most popular social network (Facebook)—this font has already been used in masked priming experiments (e.g., Duyck & Warlop, 2009; Silvia, Jones, Kelly, & Zibaie, 2011). This font has a narrow interletter spacing, thus maximizing the chances to capture the effects from multi-letter homoglyphs (if any) during printed word recognition (e.g., *documento–DOCUMENTO* vs. *docurnento–DOCUMENTO* vs. *docusnento–DOCUMENTO*). In Experiment 2, we employed Calibri. This is the default font in the most popular office package (Microsoft Office; for masked priming experiments using Calibri font, see Chen, Peltola, Ranta, & Hietanen, 2016; Tan & Yap, 2016). As Calibri has a wider interletter spacing than Tahoma (e.g., *document-DOCUMENTO vs. docurnento-DOCUMENTO vs. docusnento-DOCUMENTO*), it offers a useful scenario to test whether the effects of multi-letter homoglyphs are restricted to a special case of narrow-spaced fonts.

Experiment 1

Method

Participants. Thirty undergraduate psychology students from the Universitat de València (Spain), all of them native speakers of Spanish, took part in the experiment. All participants signed an informed consent form before starting the experiment.

Materials. The set of word stimuli was composed of 240 Spanish words extracted from the EsPal subtitle database (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013). The average Zipf frequency was 4.39 (range = 2.98–6.11), the average number of letters was 7.8 (range = 5–12), and the average OLD20 was 2.1 (range = 1– 7.8). All these words had the letters "m" or "d" in a middle position (e.g., *DOCUMENTO* [document]; *PRESIDENTE* [president]). Target words were presented in uppercase and were preceded by (a) a lowercase identity prime (identity condition; *documento–DOCUMENTO*, *presidente–PRESIDENTE*); (b) a nonword lowercase prime in which the letter 'm' (or 'd') from the base word was replaced by 'rn' (or 'cl'; homoglyph condition; *docurnento–DOCUMENTO*, *presiclente–PRESIDENTE*); or (c) a nonword prime in lowercase, in which the letter initial letter of the multi-letter homoglyph was replaced by another letter that kept the same syllabic structure (e.g., rn→sn; mean bigram token frequency

[3] Retrieved from https://msdn.microsoft.com/en-us/library/windows/desktop/dd374047(v=vs.85).aspx

[4] We acknowledge that there may be other options at creating the control condition. We employed "sn" as a control of the homoglyph "rn" because the two pairs kept the same syllabic structure and had similar bigram frequencies.

per million = 339 vs. 335, respectively, $p > .40$; control condition; *docusnento–DOCUMENTO*; *presiglente–PRESIDENTE*). We also created a set of 240 nonwords matched on letter length, transition frequencies, and subsyllabic elements with the words using Wuggy (Keuleers & Brysbaert, 2010)—the added constraint was that the letter m/d should appear in a middle position (e.g., *CLIMERO*; *VADRO*). The prime-target manipulation for the nonword targets was the same as that for word targets. Prime-target pairs were rotated across the three priming conditions in a Latin square manner, thus resulting in three lists. The complete set of words/nonwords is available at http://www.uv.es/amarhe5/glyphs.pdf.

Procedure. The experimental session took place individually in a quiet lab. A windows computer equipped with DMDX (Forster & Forster, 2003) was employed to present the stimuli and register the responses. Each trial started with a 500-ms pattern mask—a series of '#' symbols—in the center of a CRT screen for 500 ms. Then, the mask was immediately replaced by a 50-ms lowercase prime, which, in turn, was substituted by a target stimulus in uppercase. The target stimulus remained on the screen until the participant responded or a 2-s deadline had passed. Participants were instructed to press, as quickly and accurately as possible, a key labeled "sí" [yes] if the letter string formed a legitimate word in Spanish or a key labeled "no" if the letter string did not form a word. The stimuli were presented in 16-pt Tahoma. Each participant received a randomized order of trials. There were 16 practice trials before the 240 experimental trials. The session lasted for around 18 min.

Results and Discussion

Error responses were omitted from the latency analyses (2.4% for words and 6.3% for nonwords). To remove anticipatory responses, correct latencies faster than 250 ms were also excluded from the analyses (one data point; i.e., less than 0.01% of the data). The deadline for responding was 2 s, so that there could not be reaction time (RTs) longer than 2,000 ms. The averages of each of the three prime-target conditions (identity, homoglyph, control) for each dependent variable—mean correct RT and accuracy—are shown in Table 1. Words and nonwords were analyzed separately because masked form priming is typically restricted to word targets.

For the RT analyses, we employed linear mixed effects models that included prime–target relationship as a fixed factor, and subject and item as random factors. We used the maximal random structure model using the lmer package in R (Bates, Mächler, Bolker, & Walker, 2015)—LME_RT = lmer(−1,000/RT ~ primetype + [primetype + 1|item] + [primetype + 1|subject], data = RTdata)—as these models require that the underlying data follow approximately a normal distribution, RTs were transformed to decrease the positive skew of raw RTs. The three priming conditions were encoded in the model as −1, 0, +1 so that we could test the two planned comparisons: (a) homoglyph condition versus control condition, and (b) identity condition versus homoglyph condition. The lmerTest package in R (Kuznetsova, Brockhoff, & Christensen, 2016) was used to estimate the p values corresponding to the t tests. Similar analyses were conducted on the accuracy data, except for the use of generalized linear models—accuracy for each response was encoded as 1 (correct) and 0 (incorrect). For the interested readers, $F1$ and $F2$ ANOVAs yielded the same pattern of significant results as those reported here.

Table 1
Mean Lexical Decision Times (in Milliseconds) and Accuracy (in Parentheses) for Words and Nonwords in Experiment 1

	Identity	Homoglyph	Control
Words	573 (.979)	576 (.975)	595 (.973)
Nonwords	719 (.937)	716 (.935)	727 (.940)

Word data. On average, responses to target words were approximately 19 ms faster in the homoglyph condition than in the control condition ($t = 4.67$, $\beta = 0.056$, $SE = 0.012$, $p < .001$). In addition, there was only a small 3-ms nonsignificant advantage of the identity condition over the homoglyph condition ($t = -1.45$, $\beta = 0.015$, $SE = 0.013$, $p = .26$). (A post hoc analysis showed that this pattern of effects was virtually the same for the multi-letter homoglyphs rn/m and cl/d.) Accuracy was very high (0.976) and the statistical analyses did not show any significant effects (both ps $> .46$).

Nonword data. None of the effects approached significance in the latency or error data (all ps $> .20$).

Results showed faster word identification times when the prime was composed of a multi-letter homoglyph than when it was preceded by an orthographic control (i.e., *documento–DOCUMENTO* < *docusnento–DOCUMENTO*). Furthermore, the multi-letter homoglyph activated its corresponding visually similar letter to a very large degree, as deduced from the similar word identification times for the multi-letter homoglyph condition and the identity condition (i.e., *documento–DOCUMENTO* = *docusnento–DOCUMENTO*).

The question now is to what extent this pattern of data is due to specific characteristics of the font employed in the experiment. Keep in mind that Tahoma is characterized by a narrow spacing between letters. In Experiment 2, we employed the same materials and procedure as in Experiment 1 except that we employed Calibri. This font, being the default font in Microsoft Office, is currently one of the most prevalent fonts and, importantly, it has a wider interletter spacing than Tahoma. We acknowledge that another strategy could have been to increase interletter spacing in Tahoma. However, we must bear in mind that Tahoma was designed with a specific interletter spacing in mind. Finally, as the size of the effects could be smaller than those in Experiment 1, sample size was increased to 36 participants (i.e., 2,880 data points in each priming condition).

Experiment 2

Method

Participants. Thirty-six new students from the same pool as in Experiment 1 participated in the experiment.

Materials and procedure. The materials and procedure were the same as in Experiment 1, except that the font was 18-pt Calibri.

Results and Discussion

As in Experiment 1, incorrect responses (2.9% for words and 4.9% for nonwords) and RTs shorter than 250 ms (zero data points for words; four data point for nonwords [<0.05% of the data]) were excluded from the RT analyses. The mean correct RT and accuracy in each condition are shown in Table 2. The statistical tests paralleled those from Experiment 1.

Word data. Word identification times were, on average, 10 ms faster in the homoglyph condition than in the control condition ($t = 2.40$, $\beta = 0.020$, $SE = 0.008$, $p = .018$). Furthermore, the 6-ms advantage of the identity condition over the homoglyph condition was significant ($t = -2.51$, $\beta = 0.021$, $SE = 0.008$, $p = .017$). As in Experiment 1, accuracy was extremely high (0.971) and neither of the planned comparisons approached significance in the accuracy analyses (both $ps > .24$).

Nonword data. There were no signs of priming effects in the latency or error data (all $ps > .19$).

Results showed that the identity condition only produced slightly faster word identification times than the multi-letter homoglyph condition (6 ms; it was 3 ms in Experiment 1). In addition, we found an advantage of the multi-letter homoglyph priming condition over the orthographic control condition—note that it was somewhat smaller than in Experiment 1 (10 vs. 19 ms, respectively).

To examine the similarities and differences between the findings with Tahoma (Experiment 1) and Calibri (Experiment 2) fonts, we conducted a combined analysis of Experiments 1 and 2 with Experiment as a between-subjects factor. Results showed similar word response times for the identity and the multi-letter homoglyph conditions ($t = -1.35$, $\beta = 0.015$, $SE = 0.011$, $p = .18$)—this pattern was similar in the two experiments, as deduced from the lack of a significant interaction ($t = -0.40$, $p > .68$). When examining the advantage of the homoglyph condition over the control condition ($t = 5.30$, $\beta = 0.056$, $SE = 0.011$, $p < .001$), the joint analysis showed that it was greater in Experiment 1 (Tahoma) than in Experiment 2 (Calibri; $t = -2.56$, $\beta = -0.036$, $SE = 0.014$, $p = .013$). Therefore, albeit to a slightly lesser degree, Calibri font is subject to letter confusability from multi-letter homoglyphs.[5]

General Discussion

The letters that compose the words in printed Roman script are separated by small whitespaces that signal their boundaries (see Figure 1). For simplicity, models of printed word recognition inspired in the interactive activation model assume the existence of well-defined discrete slots for each letter (e.g., the visual features of "H" and "A" in HAND would be processed independently). Alternatively, the Bayesian Reader model (Norris & Kinoshita, 2012) posits that there is some degree of uncertainty at assigning visual objects to letters in the initial moments of processing. This latter assumption is consistent with anecdotal evidence that suggests that nonwords composed of multi-letter homoglyphs such as *docurnent* can be confusable with their base word. To examine whether nonwords composed of multi-letter homoglyphs such as *docurnent* activate their visually similar base words in the early stages of word processing, we conducted two masked priming experiments using two very common fonts: Tahoma (Experiment 1) and Calibri (Experiment 2)—note that Tahoma has a narrow interletter spacing. Results showed a response time advantage of the multi-letter homoglyph priming condition over the orthographic control condition in the two experiments (19 ms in Experiment 1; 10 ms in Experiment 2; i.e., *documento–DOCUMENTO* faster than *docusnento–DOCUMENTO*). Furthermore, the identity condition only showed a minimal advantage over the multi-letter homoglyph condition (a 3-ms difference in Experiment 1 and a [significant] 6 ms difference in Experiment 2). That is, in the early moments of processing, the perceptual system does not accurately perceive the whitespaces around the "r" and the "n" in *docurnento* (i.e., *docurnento* and *documento* generate a similar perceptual input). The greater effectiveness of multi-letter homoglyphs with the Tahoma than with the Calibri font probably reflects the fact that the visual features of nearby letters are closer with the Tahoma font (e.g., compare the homoglyph "rn" in docurnento [Tahoma] and docurnento [Calibri]). Taken together, these findings have important implications both at the theoretical level (i.e., how visual similarity extends across letters in models of printed word recognition) and the applied level (i.e., for Internet administrators/users and font designers).

At the theoretical level, the presence of faster responses in the multi-letter homoglyph condition than in the control condition with easily legible printed script reveals that in the early moments of word processing, the cognitive processes responsible for visual word recognition are highly resilient to potentially noisy signal (e.g., the gap between "r" and "n" in "rn"), and this is the case even when the stimuli are presented in a visually familiar format (e.g., *docurnent*). Therefore, despite the presence of visual cues (i.e., whitespaces) between letters, there is still some ambiguity at assigning the words' visual objects to letters at the early moments of processing. This phenomenon adds to the presence of uncertainty concerning letter identity and letter position during the initial moments of printed word recognition (e.g., see Gomez et al., 2008; Norris et al., 2010). The flexibility at tolerating large shape variations across letter features and letters during orthographic processing probably arises from the fact that adult readers have extensive experience with very different forms of writing (e.g., handwriting; Barnhart & Goldinger, 2010; Grainger et al., 2016; Hannagan et al., 2012). Indeed, unlike OCR engines, human readers can read distorted stimuli such as captchas (e.g., critics), partially mutilated words (e.g., hotel), or low-resolution faxes (e.g., associations) without much trouble (see Hannagan et al., 2012, for evidence of sizable masked repetition priming effects when the primes were composed of captchas; see also Perea,

Table 2
Mean Lexical Decision Times (in Milliseconds) and Accuracy (in Parentheses) for Words and Nonwords in Experiment 2

	Identity	Homoglyph	Control
Words	604 (.977)	610 (.965)	620 (.970)
Nonwords	728 (.953)	733 (.947)	726 (.953)

[5] Response times for words were, on average, 20 ms faster in Experiment 1 than in Experiment 2 ($t = 2.61$, $\beta = 0.110$, $SE = 0.042$, $p = .01$). Nonetheless, this difference should be taken with caution because this is a post hoc analysis of between-subjects data that were not collected with random assignment (i.e., we ran Experiment 1, and then Experiment 2). More important, as masked priming effects reflect a "savings" effect (see Gomez, Perea, & Ratcliff, 2013, for modeling evidence with the diffusion model), there is no theoretical reason why the overall difference in word response times would have affected the pattern of masked priming effects.

Comesaña, Soares, & Moret-Tatay, 2012, for similar evidence with mutilated prime words).

How can models of printed word recognition account for the present findings? As acknowledged by McClelland et al. (2014), the interactive activation model—for simplicity— "assumes discrete slots for letters" (p. 1181). Therefore, as it stands, this model predicts similar word identification times for *documnent–DOCUMENT* and *docusnent–DOCUMENT*, which, in turn, would be longer than the word identification times for *document–DOCUMENT*. Obviously, a similar reasoning applies to the other interactive-activation models of printed word recognition that also use the font designed by Rumelhart and Siple (1974). Nonetheless, the extension of the interactive activation model to auditory word recognition (i.e., the TRACE model) assumes "some spread of phonological features producing overlap between adjacent slots" (McClelland et al., 2014, p. 1187). If this idea were extended to the interactive activation model and its successors (i.e., using visual features instead of phonological features), this would mean that multi-letter homoglyphs could activate their visually similar letter representations, thus capturing the observed effects. However, this modification would also require a letter level more sophisticated than the Rumelhart and Siple (1974) uppercase font (see Mewhort & Johns, 1988, for criticism on the oversimplification in the letter-feature and letter levels in the coding scheme of the interactive-activation model). Importantly, the Bayesian Reader model (Norris & Kinoshita, 2012) can readily capture the observed pattern of findings because this model assumes that, in the first moments of processing, there is uncertainty when mapping visual features to letters. As Norris and Kinoshita (2012) indicated, "early in processing, there might be so much uncertainty as to how many letter objects are present that, for example, *care* might be as likely as a three-letter word" (p. 527). Nonetheless, the current version of the Bayesian Reader model needs further refinement: For simplicity, it assumes that all letters are equally confusable and it does not make any specific claims concerning the mapping of visual objects onto letters.

Clearly, the present data call for a refinement of the perceptual front end of models of printed word recognition. As Balota et al. (2006) put it, "What is the glue that puts the features together?" (p. 289) This question is related to a fundamental issue in visual perception: how the varying features from visual objects can be perceived as a whole (see Wolfe, 2012). A common view is that the "glue" that combines the letter features into letters is focal attention (i.e., conscious processing; see Treisman & Gelade, 1980, for discussion). Nonetheless, as Dehaene et al. (2004) proposed, conscious processing may not be a requirement when binding the visual components of letters or words, as this is a highly overlearned process that may involve dedicated neural pathways that combine the letter features into abstract letter units (see Keizer, Hommel, & Lamme, 2015, for a similar observation; see also Dehaene, Cohen, Sigman, & Vinckier, 2005, for a neural model of printed word recognition). Indeed, it has been claimed that one of the processing deficits of individuals with dyslexia is at binding the visual features of letters and words (see Pammer, 2014). Although an answer to the binding problem in printed word recognition would undeniably be beyond the scope of the present study, the high degree of perceptual similarity between *docurnent* and *document* at the early stages of word recognition suggest that, as occurs with other visual objects, Gestalt principle of good continuation of form also apply to letter/word recognition (i.e., rn→m; see Rosa, Perea, & Enneson, 2016, for evidence of this principle when deleting visual features from letters in printed word recognition; see also Pelli et al., 2009, for discussion of Gestalt principles in letter identification). Further research should be conducted to determine, in detail, the role of visual similarity with multi-letter homoglyphs during visual word recognition and reading. As a reviewer pointed out, one potential avenue would be to manipulate the interletter spacing of the prime stimuli using the same font (e.g., *docurnent-DOCUMENT* vs. *docur nent-DOCUMENT*)—importantly, this could combined with the recording of event-related potentials to unfold the time course of the effect. Another line of research could examine the processing of multi-letter homoglyphs in a more natural scenario: Although participants read sentences and their eyes are monitored, this could be combined with a gaze contingent boundary change paradigm (Rayner, 1975) to assess the processing of multi-letter homoglyphs in the parafovea.

At an applied level, the present data offer empirical support to the suspicion that nonwords composed of multi-letter homoglyphs such as *docurnent* are perceptually very similar to their base words. That is, a domain name such as rnicrosoft.com could be easily misread as microsoft.com (or sarnsung.com instead of samsung.com, ibrn.com instead of ibm.com, etc.). Therefore, potentially malicious imposters can buy these domain names instead of the real names to let innocent users into thinking that they are on the proper website. The result is that naïve users may give away passwords and private information. How can Internet administrators avoid these potentially threatening issues? An initial obvious solution is to buy those domain names that may be potentially confusable with the real ones. This would involve not only those domain names that employ multi-letter homoglyphs (e.g., rnicrosoft.com) but also single-letter homoglyphs (e.g., "0" and "O," as in MICR0S0FT.COM). A complementary option would be to design fonts that minimize this type of letter confusion when reading a domain name in a web browser (e.g., zero could be written as "0," as in Consolas font) together with a wide interletter spacing (e.g., r nicrosoft.com would not be easily confusable with microsoft.com).

In summary, we found that, at the early moments of word processing, nonwords created by replacing a letter with a multi-letter homoglyph (e.g., "m" with "rn," as in *docurnent*) are quite effective at activating their corresponding base words. At the theoretical level, this finding is a demonstration that there is some degree of ambiguity at mapping the visual objects that constitute the words onto letters, thus requiring more elaborated accounts of the links between the visual feature level and the letter levels in future implementations of models of printed word recognition. At an applied level, Internet users should be aware that malicious attackers might trick them with domain names that visually resemble the real websites (e.g., rnicrosoft.com), with the risk of exposing confidential information.

References

Adelman, J. S. (2011). Letters in time and retinotopic space. *Psychological Review, 118,* 570–582. http://dx.doi.org/10.1037/a0024811

Balota, D., Yap, M. J., & Cortese, M. J. (2006). Visual word recognition: The journey from features to meaning (A travel update). In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 285–375). Amsterdam, the Netherlands: Academic Press. http://dx.doi.org/10.1016/B978-012369374-7/50010-9

Barnhart, A. S., & Goldinger, S. D. (2010). Interpreting chicken-scratch: Lexical access for handwritten words. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 906–923. http://dx.doi.org/10.1037/a0019258

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Bohm, T. (2015). Letter and symbol misrecognition in highly legible typefaces for general, children, dyslexic, visually impaired and ageing readers. *Information Design Journal, 21,* 34–50. http://dx.doi.org/10.1075/idj.21.1.05boh

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences, 18,* 90–98. http://dx.doi.org/10.1016/j.tics.2013.11.005

Chen, T., Peltola, M. J., Ranta, L. J., & Hietanen, J. K. (2016). Affective priming by eye gaze stimuli: Behavioral and electrophysiological evidence. *Frontiers in Human Neuroscience, 10,* 619. http://dx.doi.org/10.3389/fnhum.2016.00619

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,* 204–256. http://dx.doi.org/10.1037/0033-295X.108.1.204

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117,* 713–758. http://dx.doi.org/10.1037/a0019738

Davis, M., & Suignard, M. (2012). *Unicode security considerations*. Unicode Technical Report No. 36. Retrieved from http://www.unicode.org/reports/tr36/

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences, 9,* 335–341. http://dx.doi.org/10.1016/j.tics.2005.05.004

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J.-B., Le Bihan, D., & Cohen, L. (2004). Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence. *Psychological Science, 15,* 307–313. http://dx.doi.org/10.1111/j.0956-7976.2004.00674.x

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods, 45,* 1246–1258. http://dx.doi.org/10.3758/s13428-013-0326-1

Duyck, W., & Warlop, N. (2009). Translation priming between the native language and a second language: New evidence from Dutch-French bilinguals. *Experimental Psychology, 56,* 173–179. http://dx.doi.org/10.1027/1618-3169.56.3.173

Finkbeiner, M., & Coltheart, M. (2009). Letter recognition: From perception to representation. *Cognitive Neuropsychology, 26,* 1–6. http://dx.doi.org/10.1080/02643290902905294

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 680–698. http://dx.doi.org/10.1037/0278-7393.10.4.680

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers, 35,* 116–124. http://dx.doi.org/10.3758/BF03195503

Gabrilovich, E., & Gontmakher, A. (2002). The homograph attack. *Communications of the ACM, 45,* 128. http://dx.doi.org/10.1145/503124.503156

Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked versus unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1731–1740. http://dx.doi.org/10.1037/a0032333

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115,* 577–600. http://dx.doi.org/10.1037/a0012667

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

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

Grainger, J., & van Heuven, W. J. B. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 1–23). Hauppauge, NY: Nova Science.

Hannagan, T., Ktori, M., Chanceaux, M., & Grainger, J. (2012). Deciphering CAPTCHAs: What a Turing test reveals about human cognition. *PLoS ONE, 7,* e32121. http://dx.doi.org/10.1371/journal.pone.0032121

Homoglyph. (n.d.). *In Wikipedia*. Retrieved on February 21, 2017, from http://en.wikipedia.org/wiki/homoglyph

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 209–229. http://dx.doi.org/10.1037/0096-1523.33.1.209

Keizer, A. W., Hommel, B., & Lamme, V. A. F. (2015). Consciousness is not necessary for visual feature binding. *Psychonomic Bulletin & Review, 22,* 453–460. http://dx.doi.org/10.3758/s13423-014-0706-2

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42,* 627–633. http://dx.doi.org/10.3758/BRM.42.3.627

Kinoshita, S., Robidoux, S., Mills, L., & Norris, D. (2014). Visual similarity effects on masked priming. *Memory & Cognition, 42,* 821–833. http://dx.doi.org/10.3758/s13421-013-0388-4

Krammer, V. (2006). Phishing defense against IDN address spoofing attacks. In *Proceedings of the 4th Annual Privacy Security Trust Conference 2006* (pp. 275–284). New York, NY: ACM Press. http://dx.doi.org/10.1145/1501434.1501473

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)*. R package Version 2.0–33. Retrieved from http://CRAN.R-project.org/package=lmerTest

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

McClelland, J. L., Mirman, D., Bolger, D. J., & Khaitan, P. (2014). Interactive activation and mutual constraint satisfaction in perception and cognition. *Cognitive Science, 38,* 1139–1189. http://dx.doi.org/10.1111/cogs.12146

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,* 375–407. http://dx.doi.org/10.1037/0033-295X.88.5.375

Mewhort, D. J. K., & Johns, E. E. (1988). Some tests of the interactive-activation model for word identification. *Psychological Research, 50,* 135–147. http://dx.doi.org/10.1007/bf00310174

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review, 113,* 327–357. http://dx.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,* 517–545. http://dx.doi.org/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,* 254–271. http://dx.doi.org/10.1016/j.jml.2009.11.002

Pammer, K. (2014). Temporal sampling in vision and the implications for dyslexia. *Frontiers in Human Neuroscience, 7,* 933. http://dx.doi.org/10.3389/fnhum.2013.00933

Pelli, D. G., 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,* 36–49. http://dx.doi.org/10.1080/13546800802550134

Perea, M., Comesaña, M., Soares, A. P., & Moret-Tatay, C. (2012). On the role of the upper part of words in lexical access: Evidence with masked priming. *The Quarterly Journal of Experimental Psychology, 65,* 911–925. http://dx.doi.org/10.1080/17470218.2011.636151

Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). R34D1NG W0RD5 W1TH NUMB3R5. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 237–241. http://dx.doi.org/10.1037/0096-1523.34.1.237

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition, 31,* 829–841. http://dx.doi.org/10.3758/BF03196438

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language, 51,* 231–246. http://dx.doi.org/10.1016/j.jml.2004.05.005

Perea, M., Marcet, A., Uixera, B., & Vergara-Martínez, M. (2017). Eye movements when reading sentences with handwritten words. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*. Advance online publication. http://dx.doi.org/10.1080/17470218.2016.1237531

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81. http://dx.doi.org/10.1016/0010-0285(75)90005-5

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. http://dx.doi.org/10.1016/j.actpsy.2016.05.016

Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review, 81,* 99–118. http://dx.doi.org/10.1037/h0036117

Silvia, P. J., Jones, H. C., Kelly, C. S., & Zibaie, A. (2011). Masked first name priming increases effort-related cardiovascular reactivity. *International Journal of Psychophysiology, 80,* 210–216. http://dx.doi.org/10.1016/j.ijpsycho.2011.03.009

Smith, R. (2007). An overview of the Tesseract OCR engine. *Proceedings of the Ninth International Conference on Document Analysis and Recognition, IEEE, 2,* 629–633.

Tan, L. C., & Yap, M. J. (2016). Are individual differences in masked repetition and semantic priming reliable? *Visual Cognition, 24,* 182–200. http://dx.doi.org/10.1080/13506285.2016.1214201

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136. http://dx.doi.org/10.1016/0010-0285(80)90005-5

Vergara-Martínez, M., Perea, M., Gómez, P., & Swaab, T. Y. (2013). ERP correlates of letter identity and letter position are modulated by lexical frequency. *Brain and Language, 125,* 11–27. http://dx.doi.org/10.1016/j.bandl.2012.12.009

von Ahn, L., Maurer, B., McMillen, C., Abraham, D., & Blum, M. (2008). reCAPTCHA: Human-based character recognition via Web security measures. *Science, 321,* 1465–1468. http://dx.doi.org/10.1126/science.1160379

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,* 221–243. http://dx.doi.org/10.3758/bf03196158

Wolfe, J. M. (2012). The binding problem lives on: Comment on Di Lollo. *Trends in Cognitive Sciences, 16,* 307–308. http://dx.doi.org/10.1016/j.tics.2012.04.013

Yakup, M., Abliz, W., Sereno, J., & Perea, M. (2015). Extending models of visual-word recognition to semicursive scripts: Evidence from masked priming in Uyghur. *Journal of Experimental Psychology: Human Perception and Performance, 41,* 1553–1562. http://dx.doi.org/10.1037/xhp0000143

Received April 25, 2017
Revision received July 13, 2017
Accepted July 19, 2017 ■