Journal of Memory and Language 68 (2013) 85–97

Contents lists available at SciVerse ScienceDirect

Journal of Memory and Language

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

The role of orthographic syllable structure in assigning letters to their position in visual word recognition

Marcus Taft *, Lidija Krebs-Lazendic

University of New South Wales, Sydney, Australia

ARTICLE INFO

Article history:
Received 11 January 2012
revision received 17 October 2012
Available online 17 November 2012

Keywords:
Lexical processing
Visual word recognition
Letter transposition
Orthographic processing
Syllable structure
Letter position

ABSTRACT

The way in which letters are assigned their position when recognizing a visually presented word was examined in three experiments using nonwords created by transposing the two medial consonants of a bisyllabic baseword (e.g., *nakpin*, *semron*). The difficulty in responding to such "TL" nonwords in a lexical decision task was shown to be lower when the medial consonants of the baseword formed a complex coda (e.g., the *rm* of *sermon*) than when they comprised a separate coda and onset (e.g., the *p* and *k* of *napkin*). The same result was shown in false positive responses to nonwords when their visibility was degraded through masking. In addition, these TL effects were just as strong for nonwords like *nakpin* as they were for nonwords whose medial consonants formed a complex coda like *warlus*, but whose baseword was syllabified between those consonants (e.g., the *l* and *r* of *walrus*). Such findings are a challenge for most current models of letter position assignment. Instead, they can be explained by an account where bisyllabic words are stored in lexical memory with a structure that maximizes the coda of the first syllable and where medial consonants are tried out in all viable subsyllabic slots.

© 2012 Elsevier Inc. All rights reserved.

Introduction

In order to recognize a visually presented word, it is necessary to ensure that its component letters are assigned to their appropriate position so that the mental representation of that word can be activated. Given the well-established finding that a word can be easily recognized even when two of its letters have been transposed, it is apparent that the assigned position for a letter need not correspond exactly to its position in the presented letter-string. This is especially true for medial letters given that their transposition (as in *nakpin* for example) is generally less disruptive to recognition of the baseword (*napkin*) than are transpositions involving initial or final letters (e.g., *anpkin* or *napkni* respectively: e.g., Bruner & O'Dowd, 1958; Chambers, 1979; Perea & Lupker, 2003; Taft & Nillsen, 2012; White, Johnson, Liversedge, & Rayner, 2008).

Different experimental paradigms have been used to demonstrate the ease of activating a word despite its transposed letters. For example, lexical decisions to a nonword are harder when it creates a real word through letter transposition (e.g., *nakpin*) relative to a nonword (such as *naktin*) that does not (e.g., Andrews, 1996; Chambers, 1979; Frankish & Turner, 2007; Lee & Taft, 2009; O'Connor & Forster, 1981; Perea & Carreiras, 2006; Perea & Lupker, 2004; Taft & Nillsen, 2012). This transposed-letter ("TL") interference effect indicates that the baseword was successfully activated even though not all of its letters were in their correct position, and such lexical activation makes it hard to classify the stimulus as a nonword. Another approach (e.g., Forster, Davis, Schoknecht, & Carter, 1987; Lupker, Perea, & Davis, 2008; Perea & Lupker, 2004; Schoonbaert & Grainger, 2004) has been to show that recognition of the (uppercase) base word (e.g., *NAPKIN*) is facilitated by the prior masked (lowercase) presentation

* Corresponding author. Address: School of Psychology, University of New South Wales, Sydney, NSW 2052, Australia. Fax: +61 2 9385 3641.
E-mail address: m.taft@unsw.edu.au (M. Taft).

0749-596X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jml.2012.10.004

of its TL version (e.g., *nakpin*) relative to the different-letter control (e.g., *naktin*).

These TL effects are important for our understanding of reading because they provide a window into the way in which information about letter position is used to make contact with stored lexical representations. There are several different approaches that have been taken in relation to this issue. First, there are models of word recognition that incorporate a sublexical level of processing whereby a word is activated via a set of "open" bigrams (e.g., Grainger & van Heuven, 2003; Grainger & Whitney, 2004; Schoonbaert & Grainger, 2004; Whitney, 2001; Whitney & Cornelissen, 2005, 2008). These are all the adjacent and non-adjacent pairs of letters found within the word in their specific order (e.g., *fl*, *fa*, *fp*, *la*, *lp*, and *ap* in *flap*). A TL nonword activates its baseword by virtue of the fact that most of their open bigrams overlap. For example, the TL nonword *falp* activates all the bigrams corresponding to the baseword *flap* except for *la*.

Davis (2010) presents an alternative spatial coding account of letter position coding (see also Davis, 2006; Davis & Bowers, 2004, 2006) where the lexical representation for each word is sensitive to a particular pattern of activation for its component letters, reflecting their relative position. So, the lexical representation for *flap* will be most responsive to a pattern of activation where *f* is activated more strongly than *l*, which is activated more strongly than *a*, which in turn is activated more strongly than *p*. When the TL nonword *falp* is presented, the level of activation for *f* and *p* is as expected for the word *flap*, while that of *a* and *l* is not very different to the expected pattern. Hence, the baseword is strongly activated and TL effects emerge.

Another approach incorporates an imprecise assignment of letters to their positional slots (e.g., Gómez, Ratcliff, & Perea, 2008; Norris, Kinoshita, & van Casteren, 2010). While the lexical representation for *flap* might be most strongly activated by the presence of *l* in the second letter-slot, it will nevertheless receive some activation from its presence in the third letter-slot, and vice versa for the letter *a*. In this way, the TL nonword *falp* will activate the lexical representation for its baseword *flap* despite the re-ordering of its medial letters.

All of the above accounts of word recognition have assumed that the only sublexical structure to play a role in orthographic processing is at the level of the single letter or bigram. However, a word is more than just a concatenation of letters or an overlapping set of bigrams. For a start, some letters function as vowels and some as consonants and there is evidence from TL manipulations that these two types of letter are processed differently in relation to the assignment of their position in the word. Perea and Lupker (2004) demonstrated stronger effects in both TL priming and TL interference for transposed consonants (e.g., *caniso* derived from the baseword *casino*) than for transposed vowels (e.g., *cisano*) and concluded that vowels must be discriminated from consonants at a very early stage of orthographic processing (cf. Berent & Perfetti, 1995). The models of orthographic processing described above have not incorporated any distinction between vowels and consonants, and Perea and Lupker (2004) make suggestions as to how they might do so.

A further important distinction can also be drawn within the category of consonants. Within a syllable, some consonants come before the vowel (i.e., are "onsets") while others come after the vowel (i.e., are "codas"). For example, *flap* has *fl* as its onset and *p* as its coda. There is considerable evidence that onsets are treated separately from the rest of the syllable in reading English (e.g., Andrews & Scarratt, 1998; Kay & Marcel, 1981; Taft, 1992; Taraban & McClelland, 1987; Treiman & Chafetz, 1987), where the rest of the syllable (or "body") is composed of the vowel and coda if there is one (e.g., the *ap* of *flap*). Lee and Taft (2009) propose a model of lexical processing that incorporates such subsyllabic structure by having the stimulus letters initially assigned to slots corresponding to their status as onset, vowel, or coda. A lexical representation (e.g., *flap*) is activated via sublexical units that separately represent onsets (e.g., *fl*) and bodies (e.g., *ap*), with the latter being activated by the letters that are assigned to the vowel and coda slots (*a* and *p* respectively). In order that the letters be assigned to their appropriate slot, consonants and vowels must be discriminated at a very early stage, as suggested by Perea and Lupker (2004). The vowel is then assigned to its corresponding slot, but the task of determining the status of the consonants as onset or coda still remains. Within each slot there are multiple positions available and these need to be assigned appropriately as well. For example, the *f* of *fl* must be placed in the first position of the onset slot and *l* in the second position of that slot.

It is relatively easy to identify which consonant should be assigned to the first position of the onset slot and which consonant should be assigned to the final position of the coda slot because the blank space before and after the letter-string makes the physical location of the first and last letters highly salient. It is the consonants that do not occur in initial or final position that are the most difficult to assign. For example, there is ambiguity as to whether the *l* of *flap* should be assigned to the second position of the onset slot or the penultimate position of the coda slot. It is only when it is assigned to the former that the lexical representation for *flap* will be activated. The same ambiguity of assignment holds for the *l* of the nonword *falp*. So, when an attempt is made to try out this *l* in the onset slot, the lexical representation for *flap* will be activated, hence generating TL effects. Because there is far less difficulty in assigning initial or final letters than medial letters to their correct slot, transpositions involving medial letters will show the stronger TL effects (e.g., Bruner & O'Dowd, 1958; Chambers, 1979; Perea & Lupker, 2003; White et al., 2008). Nevertheless, a medial transposition can often be detected, so Lee and Taft (2009) propose the existence of a supplementary mechanism that more consciously identifies the position of the consonant relative to the vowel, and assigns that consonant to its correct onset or coda slot accordingly.

Lee and Taft (2009) supported their "ambiguous assignment" account by demonstrating that TL interference does not occur for medial transpositions when the status of the medial letters as onset or coda is unambiguous. Such a situation arises with the Korean Hangul script where onset and coda position are physically identifiable even when located in the middle of a word. Because of the structure of Hangul, the words examined in that study were bisyllabic, with TL items created by exchanging the coda of the first syllable with the onset of the second. The equivalent transposition in English showed very strong TL effects (e.g., *nakpin* vs. *naktin*) in contrast to Korean, where no TL effects were observed. Lee and Taft (2009) argued that the lexical representation of each syllable of a word has its own onset/body structure ($n + ap$ and $k + in$), and TL interference arises in English when the *p* of *nakpin* is tried out as the coda of the first syllable and *k* as the onset of the second syllable, hence corresponding to the way the baseword is actually stored. Korean shows no such TL effects because assignment to the appropriate onset or coda position is defined by the physical position of the letter.

The onset-body composition of words like *napkin* is clear-cut because *pk* can be neither an onset nor a coda and, therefore, each letter in the bigram must be split between syllables when stored in lexical memory. In this way, *p* must be represented as the coda of the first syllable and *k* the onset of the second. What happens, though, when the syllable boundary is potentially ambiguous? Take the word *sermon* for example. Here, it is possible to divide the letter-string either after the *r* or after the *m* because both divisions generate a valid coda in the first syllable (*r* and *rm* respectively) and a valid onset in the second (*m* and "null" respectively). In terms of pronunciation, syllabification typically split up the medial pair of consonants (e.g., *ser-mon*), hence maximizing the onset of the second syllable in order to avoid having that syllable start with a vowel (e.g., Fallows, 1981; Pulgram, 1970; Treiman & Zukowski, 1990). However, it has been alternatively argued that, in terms of orthographic representation, polysyllabic words are divided in such a way that the longest existing coda is instantiated (e.g., *serm-on*), with this maximized first syllable being referred to as the Basic Orthographic Syllabic Structure or "BOSS" (Chen & Vaid, 2007; Rouibah & Taft, 2001; Taft, 1979, 1987, 1992, 2001, 2002; Taft, Álvarez, & Carreiras, 2007; Taft & Kougious, 2004). The BOSS is seen as optimizing the informativeness of the orthographic information contained in the first syllable of the word.

If the BOSS is a genuine unit of analysis, the orthographic structure of *sermon* will be quite different to that of *napkin* (i.e., OVCC-VC vs. OVC-OVC respectively, where V refers to a vowel, O to an onset, and C to each consonant of a coda). As a result, it might be possible to observe a difference in the magnitude of TL effects between these two types of word. In particular, the structure of the TL nonword *nakpin* is the same as that of its baseword, namely, OVC-OVC in both cases, whereas the structure of the TL nonword *semron* (also OVC-OVC) conflicts with that of its baseword where the first syllable has a complex coda and the second syllable lacks an onset (i.e., OVCC-VC). If the internal orthographic structure of words plays a role in their recognition, one might expect more confusion between a TL nonword and its baseword when their structures are congruous than when they are not. This will be examined in the present study.

In order to test such an idea more thoroughly, though, a third type of item will also be included. Here, the TL item and its baseword are also incongruous in their orthographic structure, but it is the baseword that has the OVC-OVC structure rather than the presented nonword. So, while the illegal coda *lr* has to straddle the syllable boundary in the word *walrus* (giving an OVC-OVC structure), its transposed version *rl* is an acceptable coda which means that the TL nonword *warlus* has an OVCC-VC structure according to the BOSS principle (i.e., *warl-us*). So, if compatibility of orthographic structure has an impact on the processing of letter position, items like *warlus* should be no different to items like *sermon*, since they are equally incongruent. Congruent items like *nakpin* should show stronger TL effects than both of them. For ease of discussion, the three conditions will be referred to by the baseword of these examples, namely, the "NAPKIN", "WALRUS", and "SERMON" conditions.

There is an alternative outcome that needs to be considered, however. A pair of consonants that exists as a coda (e.g., *rm* or *rl*), is very likely to be a more frequently occurring bigram than a pair that has to straddle a syllable boundary (e.g., *mr*, *lr*, *kp*, or *pk*). When a low frequency bigram is found within a letter-string, it is possible that it is initially interpreted as its transposed version if the latter is a more frequent bigram. This would be in line with the findings of Frankish and Barnes (2008) which revealed greater TL priming when bigram frequency was lower, and of Perea and Carreiras (2008) where priming was greater when the transposition created an illegal bigram than when it created a legal one (since an illegal bigram has the lowest possible frequency of occurrence, namely, zero). So, for example, the ordering of the letters *r* and *m* is far more likely to correspond to *rm* than to *mr*, and the ordering of *r* and *l* is far more likely to be *rl* than *lr*. Therefore, the likelihood that *semron* is treated as *sermon* will be greater than that of *warlus* being treated as *walrus*, hence generating stronger TL effects for the SERMON condition. This would only be the case if the compatibility of orthographic structure between the nonword and baseword were not the factor controlling their confusability, because both conditions are equally incongruent in that regard.

If bigram frequency were important, it should also lead to stronger TL effects in the SERMON condition than the NAPKIN condition, because the bigrams of the baseword for the latter (e.g., *pk*) should not be systematically different in frequency to those of the TL nonword (i.e., *kp*). Therefore, the prediction from the bigram frequency account is very different to that for the structural congruity account. Rather than the NAPKIN condition showing stronger TL effects than both the "incongruent" SERMON and WALRUS conditions, it is the SERMON condition that should produce the strongest TL effects relative to the other two conditions if bigram frequency is the controlling factor.

Finally, it should be noted that the three conditions should show the same-sized TL effects if there is no sublexical structuring of letters in the assignment of letter position (as is the case in most current theories). The same would be true if syllable structure followed the primary principle of maximal onset (e.g., Fallows, 1981; Pulgram, 1970), because the syllable boundary would always fall between the two medial consonants in both the baseword and TL nonword (giving *ser-mon*, *wal-rus* and *nap-kin*, as well as *sem-ron*, *war-lus*, and *nak-pin*).

Experiment 1

The first experiment made use of the TL interference effect. Here, the focus was on the difficulty in classifying a TL item as a nonword when briefly presented. The more readily the TL nonword activates its baseword, the greater the interference that should be observed in the lexical decision response. Manipulation of the relationship between the orthographic structure of the TL item and that of its baseword allows an examination of the role, if any, of the internal orthographic characteristics of a word in its recognition.

Method

Materials

A set of nonwords was constructed by transposing the two medial consonants of 60 two-syllable basewords. For a third of these nonwords (the NAPKIN condition), the medial consonant pair could not be used as a coda (e.g., the *kp* of *nakpin*) and this was also true for the baseword (i.e., the *pk* of *napkin*). The mean baseword frequency was 10.61 per million (SD = 29.50) according to the CELEX word frequency database (Baayen, Piepenbrock, & van Rijn, 1993), the average length 6.65 letters (SD = 0.88), and there were no single-letter-substitution neighbors (i.e., "N" = 0). For another third of the nonwords (the SERMON condition), the medial consonant pair could not be used as a coda either (e.g., the *mr* of *semron*), but here the reversal of the letters in the baseword did form a coda (i.e., the *rm* of *sermon*). Mean baseword frequency for this condition was 13.33 (SD = 29.06), mean length was also 6.65 (SD = .88), and mean N was 0.05. For the final third of the TL nonwords (the WALRUS condition), the medial consonant pair could be used as a complex coda (e.g., the *rl* of *warlus*) whereas the reversal of these letters in the baseword could not (i.e., the *lr* of *walrus*). Here, the mean baseword frequency was 13.06 (SD = 35.50), the mean length was 6.70 (SD = .86), and mean N was 0.25.

For each of these three TL conditions, every item was paired with a control nonword where the second medial consonant was replaced by a letter unrelated to the baseword. The replacement maintained the same orthographic structure as the TL item. For example, the control for *nakpin* was *naktin*, the control for *semron* was *semlon*, and the control for *warlus* was *wargus*. If there was a real word embedded in the TL nonword, this was also the case for its control (cf. the *pin* and *tin* of *nakpin* and *naktin*). All three TL conditions were closely matched to their controls in terms of N, as well as the frequency of the bigrams across the letter-string, all t's < 1. Bigram frequency (as taken from the token count of the CELEX norms[1]) was significantly higher for a TL nonword than its baseword for the SERMON condition, $t(19) = 4.85$, $p < .001$, and significantly lower for the WALRUS items, $t(19) = 3.59$, $p < .001$. The NAPKIN items did not differ from their baseword in terms of bigram frequency, $t < 1$. The items can be found in the Appendix.

A Latin-Square design was used so that no participant was exposed to both the TL nonword and its matched control. Thus, two sublists were formed with 10 items in each of the six conditions. The same set of real-word distractors was used in both sublists, consisting of 60 words of similar length and structure to the nonword targets (e.g., *medley*, *chimney*, *hostel*). There were also 12 practice items composed of words and nonwords of a similar structure to those used in the experiment.

Procedure

The lexical decision task was administered using the DMDX computer display program (Forster & Forster, 2003). Each item was presented in lowercase for 250 ms in 20 point Arial font with a 2 s delay between the response and the next item being presented. The 120 items were presented in a different random order for each participant. Participants were instructed to press the right shift key if the presented letter-string was a real English word, or the left shift key if it was not. They were asked to respond as quickly, but as accurately as possible.

Participants

The participants were 32 undergraduate psychology students from the University of New South who received course credit for participating. All were monolingual English speakers with normal or corrected-to-normal vision. They were equally allocated to one of the two sublists in a random fashion. One participant was eliminated from the experiment as a result of making more than 50% errors throughout the experiment, indicating a possible failure to attend to the task properly.

Results

The mean RTs can be found in Fig. 1a and error rates in Fig. 1b. The data were analyzed using Linear Mixed Effects (LME) modeling in R (Baayen, Davidson, & Bates, 2008).

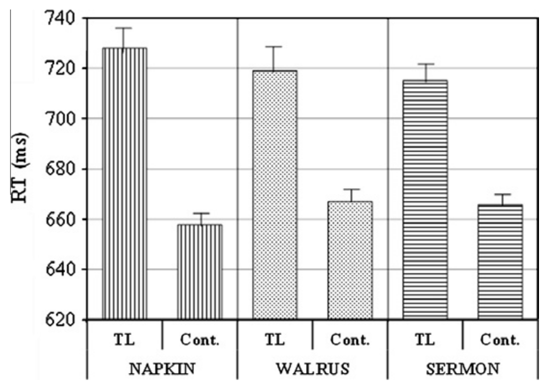

Fig. 1a. Reaction time in milliseconds for the NAPKIN, WALRUS, and SERMON conditions of Experiment 1 based on item means. Bars represent standard error.

[1] The comparison of bigrams gave the same outcome regardless of the particular measure used (e.g., log token frequency, type frequency, etc.).

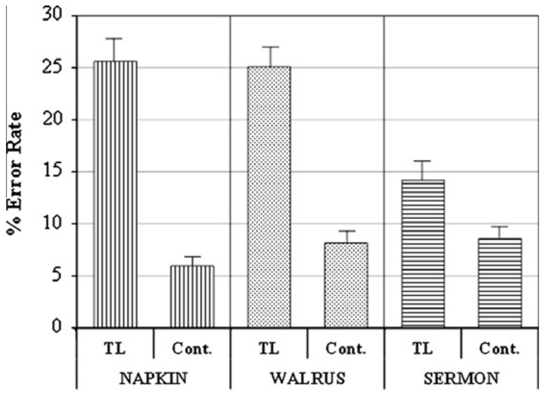

Fig. 1b. Percentage error rates for the NAPKIN, WALRUS, and SERMON conditions of Experiment 1 based on item means. Bars represent standard error.

Before fitting the model to the data, RTs faster than 200 ms and slower than 2000 ms were excluded (Baayen & Milin, 2010). A further seven outlying data points (from five different participants) were removed manually via separate inspection of RTs for each subject using quintile-quintile plots and applying the qqmath() function (Baayen, 2008).

A sequence of models was then compared across a gradually decreasing complexity of random effects structure (Baayen et al., 2008) with log RT as the dependent variable. The RTs were logarithmically transformed in order to eliminate the skewness of their distribution. The factors included in model testing were (a) a model with item intercept, subject intercept, and by factor random slopes for item as random factors, (b) a model with item intercept, subject intercept, and by factor random slopes for subject as random factors, (c) a model with item intercept and subject intercept as random factors, (d) a model with subject intercept as random factor, and (e) a model with item intercept as random factor. The likelihood ratio showed that the model with subject intercepts as random effects was the best fit to this data set (model d). Thus, this model was used for the analysis of fixed effects. For the analysis of error rates, the same model with the function for binomial data was used, with z scores being generated.

The analysis of RTs revealed a main effect of transposition (TL vs. Control), $t = 3.82$, $p < .001$, but no interaction with TL condition, $t < 1$. The analysis of error rates also showed a main effect of TL interference, $z = 8.225$, $p < .001$, but on this measure there was a significant interaction with TL condition, $z = 2.969$, $p < .01$, which arose from weaker TL interference for the SERMON condition compared to the other two conditions. Despite this, the interference for the SERMON condition was still significant on its own, $z = 2.131$, $p < .05$. There was no difference in the degree of interference between the NAPKIN and WALRUS conditions, $z = -0.108$, $p > .1$.

Discussion

It can be seen from these results that, while all TL conditions showed interference relative to their controls, there was clear evidence in the error rate for weaker interference for the SERMON condition than for the other two conditions, which did not differ from each other. Such an outcome suggests that letter-strings are treated as more than just a concatenation of letters, and that their internal orthographic structure plays some role in their processing. However, the relative weakness of the effect for the SERMON condition was not predicted. While an account in terms of congruity between the structure of the TL nonword and its baseword can explain the weaker TL effect for the SERMON condition relative to the NAPKIN condition, it cannot explain why it was also weaker than the effect for the WALRUS condition. The latter had just as incompatible an orthographic structure as the SERMON condition. An alternative account given in terms of the relative bigram frequency of the transposed consonants in the stimulus and baseword correctly predicts no difference between the NAPKIN and WALRUS conditions, but predicts greater TL interference for the SERMON condition rather than the weaker interference that was obtained. It is apparent, then, that neither account is viable.

Before considering how to explain the observed pattern of results, however, consideration needs to be given to the fact that the interaction between conditions was only found on the accuracy measure. When error rates are high, the RT measure becomes less reliable because it is based on a relatively low number of correct responses. In fact, the lack of an interaction on RTs is consistent with other lexical decision experiments that have examined whether the nature of the stimulus has an impact on the amount of TL interference observed. For example, Perea and Lupker (2004) found no interaction on their RT measure between the amount of TL interference and the status of the transposed letters as vowel or consonant. As in the present experiment, the effect only emerged on the accuracy measure. The same was true in the O'Connor and Forster (1981) study looking at the impact on TL interference of the frequency of the baseword.

Frankish and Turner (2007) ignored RTs altogether in their lexical decision experiment when finding a stronger TL effect for unpronounceable nonwords than pronounceable ones. They focused instead on the rate of false positive responses under highly degraded presentation conditions. While Experiment 1 of the present study used a shorter duration (i.e., 250 ms) than is usually adopted in lexical decision experiments (i.e., 500 ms), the stimuli used by Frankish and Turner (2007) were only of 40 ms duration and were immediately masked by a display of letter fragments. Under such degraded conditions, the error rates on TL nonwords were so high (over 50% in some conditions) that RTs would have been hard to interpret. Given that Frankish and Turner (2007) did reveal differences in accuracy between different types of TL nonword, Experiment 2 was set up using experimental conditions similar to theirs in order to provide converging evidence for the weaker interference for SERMON items relative to NAPKIN and WALRUS items. Under such degraded conditions, positional information should be particularly ambiguous and, therefore, any effects of orthographic structural constraints should be clearly observed if they occur.

Experiment 2

The results of Experiment 1 suggest that the TL effect is stronger when the baseword has a clear-cut syllable boundary between its two medial consonants (e.g., *napkin*, *walrus*) than when it has a complex coda in the first syllable as defined by the BOSS (e.g., *sermon*). If such a result is genuine, the same pattern should be observed for false positive responses when the stimuli are barely discernible. A confirmation of the results of Experiment 1 would be seen if the difference in confusion with the baseword between TL nonwords and their controls was smaller in the SERMON condition than in the other two conditions.

Method

Materials and procedure

The materials were identical to those used in Experiment 1.

A plus sign was presented as a fixation point immediately prior to the brief presentation of the target, which was followed by a 500 ms pattern mask. Instead of the mask being a display of letter fragments as used by Frankish and Turner (2007), it consisted of a row of hash-marks as favored in previous research that has examined lexical processing under masked conditions (e.g., Forster & Davis, 1984). Another change from Frankish and Turner (2007) was that the target was presented for 80 ms rather than 40 ms. The reason for this was that a pilot study with a 40 ms exposure duration established that the target could be barely seen at all. Perhaps the longer items used here compared to the five-letter strings of the Frankish and Turner (2007) study, and/or the use of hash-marks as the mask, made it harder to see the targets. The exposure duration of 80 ms was chosen because, while the target was still very hard to discern, it seemed long enough to avoid a floor effect and, hence, had the potential to reveal a sensitivity to any differences that might exist between conditions.[2]

The instructions for lexical decision were the same as in Experiment 1, except that the participants were told that the targets would be presented extremely briefly.

Participants

A new set of 38 undergraduate students was drawn from the same pool as in Experiment 1. These participants were equally allocated to each of the two sublists in a random fashion.

Results

The error rates (i.e., false positives) for the three conditions are presented in Fig. 2. As in the Frankish and Turner (2007) study, RTs will not be reported because the error rates were too high to allow reliable measurement of the latencies for correct responses (with up to 100% errors for some of the TL items).

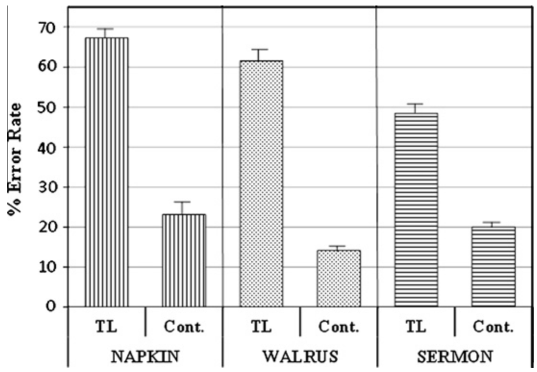

Fig. 2. Percentage error rates (false positives) for the NAPKIN, WALRUS, and SERMON conditions of Experiment 2 based on item means. Bars represent standard error.

Within the LME procedure, the likelihood ratio showed that the model with subjects and items as random effects provided the best fit. Thus, this model was used with the function for binomial data implemented and *z*-scores generated. A main effect of TL interference was revealed, $z = 10.458$, $p < .001$, but more importantly, there was a significant interaction with condition arising again from weaker TL interference for the SERMON condition when compared to the NAPKIN and the WALRUS conditions, $z = 2.285$, $p < .05$. The SERMON condition alone nevertheless produced significant TL interference, $z = 5.17$, $p < .001$. The magnitude of TL interference was the same for the NAPKIN and WALRUS conditions, with no significant interaction, $z = 1.078$, $p > .05$.

Discussion

The results obtained under highly degraded presentation conditions clearly replicate the error data of Experiment 1. While even the control condition produced error rates of over 20%, there was a highly significant TL effect, with considerable difficulty being observed in distinguishing a TL nonword from its baseword. Importantly, though, this difficulty was again significantly greater for the NAPKIN and WALRUS conditions than for the SERMON condition.

The accuracy rate for the TL items was much lower in this experiment than in Experiment 1 where the target duration was 250 ms. This is consistent with the idea put forward by Lee and Taft (2009) that the automatic processing of letter position that leads to the TL effect is supplemented by a more analytic strategy that can be applied to the letter-string. When the letter-string is presented for only 80 ms and masked, this slower, more conscious mechanism is unavailable to counter the acceptance of the TL nonword as its baseword, and a very high error rate ensues.

While the results of this experiment once more support the idea that the orthographic structure of the baseword has an impact on TL confusability, it needs to be established that the consistent reduction in TL effects for the SERMON items of Experiments 1 and 2 did not arise from

[2] Perea, Rosa, and Gómez (2005) tested masked targets with a duration of 150 ms and observed an error rate for TL nonwords that was somewhat higher than that for Experiment 1 here, but lower than that observed by Frankish and Turner (2007). The shorter target duration of 80 ms was adopted here in order to optimize the change from the presentation conditions of Experiment 1.

some characteristic of those nonwords that unwittingly made them less wordlike than the other TL items. Therefore, in order to determine the generalizability of the weaker TL interference effect for SERMON items, Experiment 3 was carried out with a new and enlarged set of materials.

Experiment 3

The purpose of this experiment was to confirm the reduced confusability between a TL nonword and its baseword when the latter had an OVCC-VC structure (e.g., *sermon*) than when it had an OVC-OVC structure (e.g., *napkin*, *walrus*). A new set of items was tested and the false positive rate was again examined under highly degraded display conditions.

Method

Materials and procedure

The design of the three TL conditions was the same as in Experiments 1 and 2. This time, though, there were 30 items in each of the three conditions, all of which had a mean length of 7.1 letters (ranging from 5 to 10). Baseword frequency was matched across the conditions according to the CELEX database (Baayen et al., 1993), with mean frequencies per million of 9.45 (SD = 26.32), 5.84 (SD = 7.55), and 8.54 (SD = 17.52) for the NAPKIN, WALRUS, and SERMON conditions respectively, all t's < 1. None of the NAPKIN or SERMON items had any single-letter-substitution neighbors, while the mean N of the WALRUS condition was 0.23. Owing to the relative shortage of English words with an OVC-OVC structure, four of the NAPKIN items and four of the WALRUS items were re-used from the previous set of materials, selected to allow suitable matching on length and baseword frequency between conditions. All 30 of the SERMON items were new.

Another change came in the design of the Control items. In previous experiments that have examined the TL effect, the control nonwords have typically been generated by substituting another letter for each of the transposed letters of the TL nonword, rather than just the second transposed letter (as in Experiments 1 and 2). The more standard approach of double substitution was adopted in Experiment 3. So, for example, the control for *meldey* (a WALRUS item) was *merney*. As in the previous experiments, the orthographic structure of each TL condition was maintained for its control, and N was matched between each TL and Control condition, t's < 1. As before, the mean bigram frequency for the SERMON items was significantly higher than for their basewords, $t(29) = 6.78$, $p < .001$, while the reverse was true for the WALRUS items, $t(29) = 2.96$, $p < .01$. The NAPKIN items again did not differ from their basewords in this regard, $t < 1$. The items used in Experiment 3 can be found in the Appendix.

Within the Latin-Square design, there were two sublists with 15 items in each of the six conditions. The same set of real-word distractors was used in both sublists, consisting of 75 words of similar length and structure to the nonword targets (e.g., *kidney*, *reptile*, *martyr*). Many of these were the baseword of TL items used in Experiments 1 and 2. There were again 12 practice items composed of words and nonwords of a similar structure to those used in the experiment.

Exactly the same procedure was followed in Experiment 3 as in Experiment 2.

Participants

A further 32 participants were drawn from the same undergraduate pool as in the other experiments. They were equally divided between each of the two sublists in a random fashion.

Results

The error rates for the three conditions are presented in Fig. 3.

Data analysis followed the same procedure as in Experiment 2. The analysis of error rates used the model with subjects and items as random effects. The function for binomial data was implemented and z-scores generated. The analysis revealed exactly the same pattern of results as observed in Experiment 2. There was a main effect of TL interference, $z = 16.1999$, $p < .001$, which was significantly reduced in the SERMON condition when compared to the WALRUS and NAPKIN conditions, $z = -2.093$, $p < .005$, though still significant for the SERMON condition on its own, $z = 6.879$, $p < .001$. Again, the WALRUS and NAPKIN conditions showed the same magnitude of TL interference, $z = 0.349$, $p > .05$.

Discussion

It appears that the reduction in the amount of TL interference observed in the first two experiments for SERMON items was generalizable to a new (and larger) set of items designed in exactly the same way. SERMON nonwords have the same OVC-OVC structure as NAPKIN nonwords (e.g., *semron* and *nakpin*), but according to the maximal coda principle of the BOSS, they differ in the structure of their baseword (i.e., *sermon* is OVCC-VC, while *napkin* is still OVC-OVC). In contrast, WALRUS nonwords have a different structure to NAPKIN nonwords (e.g., *warlus* is OVCC-VC), but their basewords have the same structure (i.e., OVC-OVC). So, given that the TL

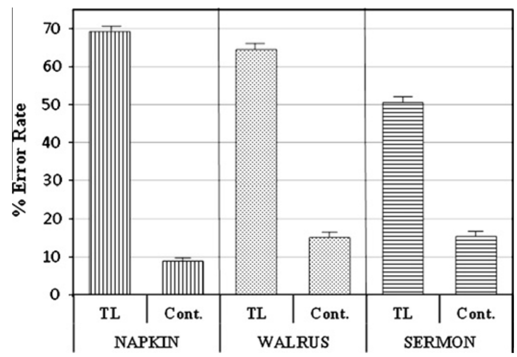

Fig. 3. Percentage error rates (false positives) for the NAPKIN, WALRUS, and SERMON conditions of Experiment 3 based on item means. Bars represent standard error.

interference for the NAPKIN condition was the same as for the WALRUS condition, but greater than for the SERMON condition, it is apparent that it is the nature of the baseword that has an impact, not the structure of the TL nonword itself. Such a sensitivity to the internal orthographic structure of the baseword has never been shown before and is incompatible with most previous accounts of letter position processing as they currently stand (e.g., Davis, 2010; Gómez et al., 2008; Grainger & van Heuven, 2003; Whitney & Cornelissen, 2008). An explanation for the findings will be considered below.

General discussion

The three experiments reported here indicate that the internal orthographic structure of a word plays a role in its recognition. When two letters of a word are transposed to create a nonword, that word is still activated as can be seen from the disruption to the accuracy of nonword classifications both when the letter-string can be clearly seen (Experiment 1) and when it cannot (Experiments 2 and 3). However, the degree to which such activation occurs appears to be constrained by the orthographic structure of the baseword. When the baseword has two medial consonants that can form a complex coda (e.g., the *rm* of *sermon*, the *nd* of *cinder*, or the *ft* of *hefty*), transposition of those consonants is more disruptive to the recognition of the word than when the two medial consonants do not form a coda (e.g., the *pk* of *napkin*, the *lr* of *walrus*, or the *tl* of *atlas*). It is apparent, then, that these two types of word are represented differently in lexical memory, which can be captured by the idea that the coda of the first syllable is maximized in the lexical representation (giving *serm-on*). This corresponds to the representation assumed by the BOSS proposal (e.g., Taft, 1979, 1987, 1992, 2001, 2002) where the informativeness of the first orthographic unit is optimized.

The BOSS is defined as an orthographic unit, but it is conceivable that maximization of the coda can be alternatively explained in phonological terms. Linguists have argued that the primary principle for syllabifying a bisyllabic word is to maximize the legal onset of the second syllable (see e.g., Fallows, 1981; Pulgram, 1970), which would place the syllable boundary between the two consonants for all three conditions in the present research (e.g., *ser-mon*, *nap-kin*, or *wal-rus*). Clearly, this is incompatible with the data that was obtained here. However, other principles might be at work, as indicated by Treiman and Zukowski (1990), such as the maintenance of decreasing sonority in the coda (referring to the amount of resonance during articulation of the phoneme). Liquids are more sonorant than nasals which are, in turn, more sonorant than obstruents (fricatives and plosives). Therefore, the nasal *m* might be grouped with the liquid *r* in *sermon* because it provides the decreasing sonority gradient that is absent in *napkin* (where *k* and *p* are both obstruents) or *walrus* (i.e., where *l* and *r* are both liquids[3]). When Treiman and Zukowski (1990) directly tested preferred syllabications, around 10% of participants placed the second medial consonant of SERMON-type items in the first syllable, though most of these participants also placed it in the second syllable, making it ambisyllabic. This never happened for NAPKIN-type or WALRUS-type items. Therefore, there is some indication that even a phonologically defined syllable might at least sometimes follow the principle of a maximal coda. However, as Treiman and Zukowski (1990) suggest, the ambisyllabicity that was shown might well have arisen from a re-syllabification of the stored representation, particularly since the spoken syllable boundary seems to depend on the speed with which the word is uttered. Hence, the preference by some people to assign the *m* of *sermon* to the first syllable may not indicate that they actually have the word stored with such a structure. Indeed, if it is phonology that determines the syllabic representation relevant to orthographic processing, one can ask how a word like *cupboard* (a NAPKIN item in Experiment 1) is able to access its representation when its orthographic analysis as *cup* and *board* is entirely incompatible with its phonological form.

Regardless of whether the maximization of the coda is based on optimizing the informativeness of the first orthographic syllable or on some phonological principle, it provides the most logical locus for the reduction in the size of the TL effect for SERMON items since that is where they differ from the other items. The main alternative to consider would seem to be one where bigram frequency plays a role in processing (e.g., Adams, 1990; Seidenberg, 1987). Consonants that form a complex coda (e.g., *rm*, *rl*) are likely to be a more frequent combination than those that do not form a coda (e.g., *mr*, *lr*), and the way in which a word is activated in lexical memory might be sensitive to such bigram frequencies. The problem with this idea, however, is that it is hard to see how it could lead to weaker TL effects for SERMON items than WALRUS items. If the letters *m* and *r* are more likely to combine as *rm* than *mr*, it should mean that *semron* will be readily interpreted as *sermon*, yet this was where the weakest TL effects were observed. It might then be argued that *rm* is so strongly entrenched in the lexical system that nothing else will be confused with it. However, the nonword *warlus*, with the more common ordering of *r* and *l*, was readily interpreted as the word *walrus* despite the *lr* being so weakly represented in the lexical system. It is hard to conceive of a model that would allow a stronger bigram to be interpreted as a weaker one. Indeed, Frankish and Barnes (2008) and Perea and Carreiras (2008) found that nonwords with low bigram frequency produced greater TL priming than those with high bigram frequency, and this supports the idea that, if bigram frequency were to have played a role in the present study, it should have revealed larger TL effects for SERMON items than WALRUS items rather than the other way around.

Consideration also needs to be given to the possibility that the results arose from confounds in the materials. Because of the difficulties in finding words in English with medial consonant pairs that cannot occur within the same syllable (as in the NAPKIN and WALRUS conditions), it was necessary to include some polymorphemic words in those conditions, most notably compounds such as *cowboy*, *rainbow*, *headline*, and *topmost*. While there were also a couple

[3] All the other WALRUS items actually had medial consonants that showed an increasing sonority gradient such as *atlas*, where *l* is more sonorant than *t*.

of compounds in the SERMON condition (i.e., *heartache*, *handout*), the transposition was intra-morphemic in these cases (e.g., *heatrache*) rather than cross-morphemic as in the other two conditions (e.g., *cobwoy*, *healdine*). Findings have been reported where morphological structure has an impact on the magnitude of the TL effect (e.g., Christianson, Johnson, & Rayner, 2005; Duñabeitia, Perea, & Carreiras, 2007; but see Rueckl & Rimzhim, 2011), so a confound between conditions on this factor needs to be taken into consideration. It is apparent, however, that morphological factors cannot account for the effects observed in the present study. When morphology has been reported to have had an impact on TL responses, it has taken the form of weaker effects for cross-morphemic transpositions than intra-morphemic ones. For example, Christianson et al. (2005) reported weaker priming effects when the transposition crossed the two morphemes of a compound word (e.g., *sunhsine*, *cobwoy*) than when it did not (e.g., *susnhine*, *heatrache*). So, according to this finding, the TL effects for the SERMON condition should have been stronger than those for the other two conditions if morphology had have played a role. Yet, they were weaker. Thus, we can dismiss morphological structure as the basis for the present results.

Another possible confound to consider is a difference amongst conditions in the phonological relationship between the TL nonword and its baseword. The most common way to create an existing coda out of two medial consonants is to use a liquid (*r* or *l*) or a nasal (*n* or *m*) in the first position. Therefore, most of the WALRUS and SERMON items included one of these consonants in order that the medial consonants formed an existing coda either in the baseword (i.e., the SERMON condition) or in the TL nonword (i.e., the WALRUS condition). In contrast, the requirement that a TL nonword and its basewords in the NAPKIN condition could not include a medial coda was met by avoiding liquid consonants (i.e., *r* and *l*). One way in which this systematic difference between conditions might have had an effect is that liquid consonants, especially *r*, generate a different pronunciation when used within a coda and when used as an onset. For example, while onset *r* is pronounced /r/ in *semron* and *makret*, not only is the *r* unpronounced in *sermon* and *market* when the participant speaks a non-rhotic dialect such as the Australian English of the present participants, but it modifies the pronunciation of the preceding vowel (giving /ɜ:/ rather than /ɛ/, or /ɑ:/ rather than /æ/). Therefore, it might be argued that the reason for weak TL effects in the SERMON condition was actually the incompatibility between the pronunciation of the TL nonword and its baseword. Indeed, Frankish and Turner (2007) and Frankish and Barnes (2008) have argued that phonology plays a role in the TL effect (though see Perea & Carreiras, 2008). The problem with this argument, however, is that the TL nonwords in the WALRUS condition had the same type of phonological relationships with their corresponding basewords as in the SERMON condition, but simply reversed. The phonemes of *walrus*, *fabric*, and *vibrate*, for example, also conflict with those of *walrus*, *farbic*, and *virbate*. So, while a phonological explanation might be argued in relation to the SERMON vs. NAPKIN results, it cannot explain why the results for the WALRUS condition were the same as for the NAPKIN condition rather than the SERMON condition.

It seems, then, that the relative weakness of TL effects for SERMON items cannot be ascribed to confounds in the materials, but rather, to an explanation that draws upon the nature of the orthographic structure that was manipulated between conditions. Such an explanation seems to require the idea that the principle of maximal coda is applied to the lexical representation accessed in reading, with complex codas being formed where possible (in accordance with principles of the BOSS). However, the explanation given cannot be based on the incompatibility between the orthographic structures of the TL nonword and its baseword, because *warlus* and *walrus* are just as incongruent in structure as are *semron* and *sermon*. So, what account can be given?

The model of Lee and Taft (2009) offers a possible way in which to incorporate the present findings. This is illustrated in Figs. 4a and 4b, where the representation for the words *napkin* and *sermon* are respectively shown, with the stimulus being a TL version of the baseword. Sublexical units exist for the onset and body which are responsive to letters appearing in particular slots corresponding to onset, vowel, and coda for each syllable. These are represented in the figures by O, V, and C respectively, with a subscript referring to the syllable. Body units are responsive to letters in the vowel slot combined with letters in the relevant coda slot. Each slot needs to cater for more than one letter (required for when the word is *strength* or *beauty*, for example), providing information about the order of those letters. For the sake of simplicity, however, the figures only depict two positions for each slot, and these are labeled "1" and "2". The first position of an onset slot and the second position of a coda slot correspond respectively to the initial and final consonant of a syllable.

When a letter-string is presented, vowels are identified and assigned to the vowel slot for the appropriate syllable. If the vowel is a single letter, it is placed in the first position of the vowel slot. Because of the space between words, the first letter of the string is readily located and, if it is a consonant, is unequivocally assigned to the first position of the first onset. The same is true for the final letter in relation to the final position of the final coda. Assignment of the other consonants to a slot is far less precise and all possibilities are tried out. This is illustrated in the figures by means of broken lines for one of the medial consonants in each of the TL nonwords (i.e., *p* in the case of *nakpin* and *r* in the case of *semron*). The same thing holds for the other medial consonant, but is not illustrated.

Successful assignment of a medial consonant is primarily determined by whether it leads to the activation of lexical information. For example, when the *p* of *nakpin* is tested in the second position of C_1 (indicated by the darker broken line), it will fully activate the body *ap* which, along with the activation of the onset *n*, will fully activate the syllabic unit *nap*. Similarly, placing *k* in the first position of O_2 will fully activate the syllabic unit *kin*. Hence the lexical representation for *napkin* will be strongly activated. There will be other units activated in competition, but never more strongly than the full activation of *nap* and

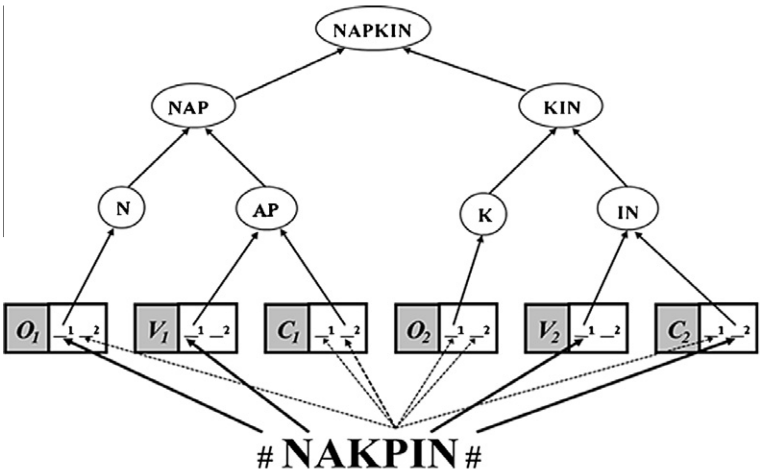

Fig. 4a. The way in which the word *napkin* would be represented in a model of lexical memory which has units corresponding to syllables, onsets, and bodies. Slots exist for onsets (O), vowels (V), and codas (C) with body units being activated when the latter two are filled appropriately. The figure shows the slot assignment for the vowels and the first and last consonants of the TL nonword *nakpin*, as well as the available choices for the letter *p* (shown as broken lines). When letter *p* fills the outer coda slot of the first syllable (shown by the darker broken line), it contributes full activation to the body unit *ap*.

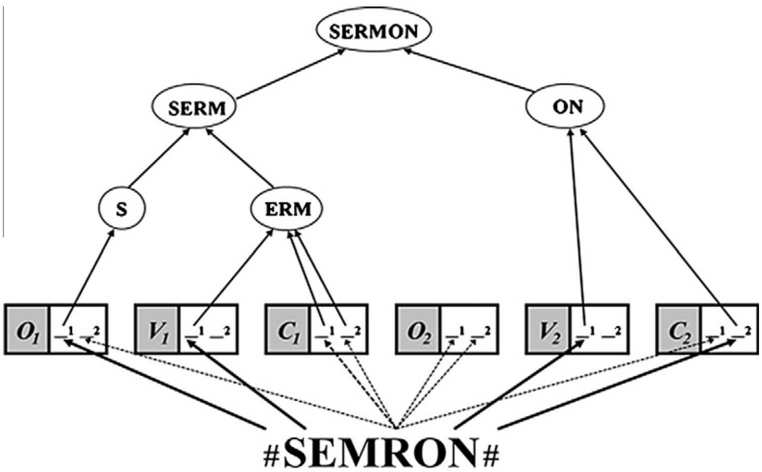

Fig. 4b. The way in which the word *sermon* would be represented by the same model. The syllabic analysis maximizes the coda of the first syllable. Assignment of the letter *r* of the TL nonword *semron* to any of the available slots (shown as broken lines) will not contribute full activation to the body unit *erm*, though assignment to the first coda slot of the first syllable (shown by the darker broken line) will partially activate that body.

kin (e.g., trying out the *k* of *nakpin* in the second position of C_1 will activate the unit *nak* that, in turn, will activate the word *naked*, but only weakly because there is no support from the unit *ed*). It is this testing-out procedure for medial consonants that can potentially explain how it is possible for *semron* to show less of a TL effect than *nakpin* (or *walrus*), as will now be explained.

When there is a single consonant in the coda of the baseword (e.g., *nap* or *wal*) or in the onset (as in *kin* or *rus*), the syllabic unit will be fully activated when that consonant is tried out in the relevant position. However, when the body of the first syllable has a complex coda (e.g., the *erm* of *serm*), each consonant will only partially activate that syllabic unit and, moreover, will potentially activate competitors more strongly. For example, when the *r* of *semron* is tried out in the first position of C_1, it will only partially activate the body *erm*, but when tried out in the second position of C_1, it will fully activate the inappropriate body *er* (cf. *serious*). Similarly, the *m* will only partially activate the body *erm* when tried out in the second position of C_1, but will fully activate its competitor *em* (cf. *seminar*). As in the NAPKIN and WALRUS conditions, there could also be other competitors when *semron* is presented such as words containing the syllable *ron* (e.g., *squadron*) when the *r* is tried out in the first position of O_1. However, the important point about the SERMON condition is that the complex coda is activated less fully by an individual letter than is a simple coda, which means that it takes longer for the units of the baseword to be fully activated. In this way, the TL item is less likely than the other conditions to be classified as the baseword when the reader makes their lexical decision response too hastily and therefore makes an error (as in Experiment 1) or has to base their judgement on information that is presented too quickly (as in Experiments 2 and 3).

The above account also fits neatly with the results of the only other reported study that has examined the impact of syllabic structure on letter assignment. Perea and Carreiras (2006) found that TL interference was just as strong in Spanish when syllables were transposed (e.g., *privemara* from the baseword *primavera*) as when non-syllabic bigrams were transposed (e.g., *primerava*). Syllables were defined in terms of the principle of maximal onset in accordance with pronunciation (i.e., *pri.me.ve.ra*) rather than in accordance with the principle of maximal coda as supported here for English (i.e., *prim.av.er.a*). However, regardless of which is appropriate for Spanish, the present account predicts the equivalent interference that was observed because the medial consonants will be tried out in all possible slots, both onset and coda, and this will activate the baseword, especially if the vowels are also tried out in the various vowel slots.

While the ambiguous-assignment account has the potential to explain the reduction in TL interference for SERMON items, however, it is weakened by the fact that it still requires further assumptions. In particular, when the *r* of *semron* is assigned to the first C_1 position and the *m* to the second C_1 position, this should activate the baseword *sermon* and putatively generate just as strong a TL effect as the items whose baseword has a single-consonant coda. There are several ways of refining the account to handle this idea. For example, it may be that letters are tried out sequentially, letter-by-letter, in each position so that when candidates are activated on the basis of one of the medial consonants (e.g., *seminar, serious*, etc.), this inhibits the attempt to assign the other consonant to the same coda slot. Alternatively, when the baseword has a single-consonant coda (e.g., *napkin*, *walrus*), assignment of the medial consonant to any of the positions in the coda slot will be effective, whereas a baseword with a double-consonant coda (e.g., *sermon*) will only be activated when each consonant is assigned to its specified position within that coda slot.

Despite the fact that the ambiguous-assignment approach undoubtedly needs refinement, it is hard to see how any of the other accounts of letter position processing (e.g., Davis, 2010; Gómez et al., 2008; Grainger & van Heuven, 2003; Whitney & Cornelissen, 2008) could even be modified to accommodate the present results while maintaining their distinctive characteristics. They would all need the inclusion of a level of orthographic representation that incorporates syllabic structure following the principle of maximal coda, but this alone cannot capture the present results. The way in which that structure is brought into play during letter assignment must be considered.

It is not obvious how an open bigram model (e.g., Grainger & van Heuven, 2003; Grainger & Whitney, 2004; Schoonbaert & Grainger, 2004; Whitney, 2001; Whitney & Cornelissen, 2005, 2008) could be set up in such a way that open bigrams extracted from the stimulus are matched to word representations with an OVCC-VC structure differently to those with an OVC-OVC structure. An argument might be made that open bigrams are more weakly formed across a syllable boundary than within a syllable. However, TL interference would then be weaker for the NAPKIN condition because the transposition is always inter-syllabic regardless of whether the syllable boundary constraint applies to the processing of the stimulus or the baseword. Neither can the spatial coding model (Davis, 2006, 2010; Davis & Bowers, 2004, 2006) readily handle the reported pattern of data because it is hard to envisage how the simple matching of relative activation levels derived from the TL nonword could be influenced by the way in which syllabic structure might be represented for the baseword.

The models based on an imprecise assignment of letters to their positional slots (e.g., Gómez et al., 2008; Norris et al., 2010) probably have the greatest potential for explaining the present results. What they would need to assume is that orthographic subsyllabic structure is represented in a similar way to Figs. 4a and 4b, and that the perceptually-based gradation of activation feeds into this structure in such a way that OVC-OVC representations are more strongly activated than OVCC-VC ones. An explanation would be needed for how the spatially determined pattern of letter activation maps onto the functionally structured lexical representation and, once the appropriate adaptations are made, it is likely to follow quite similar principles to the approach presented here.

Conclusions

The research reported here indicates that differences in the orthographic structure of a word can have an impact on its confusability with a nonword generated from the transposition of two of its letters. The difference in orthographic structure being referred to is specifically where a principle of maximal coda is applied to the syllabification rather than a principle of maximal onset that corresponds more closely to the surface pronunciation of the word. When the transposed letters occur within a complex coda (e.g., *semron* being generated by the transposition of the complex coda *rm* in *sermon*), confusion with the baseword is weaker than when the transposition exchanges a coda for an adjacent onset (e.g., *nakpin* from *napkin*, or *warlus* from *walrus*). The structure of the TL nonword itself appears to have no impact on responses.

The explanation given for this finding is couched in terms of the existence of sublexical representations of onsets and bodies (with maximized codas), and the imprecise assignment of incoming letters to slots corresponding to onsets, vowels, and codas. Previously favored models of letter-position processing have defined the similarity of a TL item and its baseword solely at the level of individual letters or bigrams. Internal orthographic structure plays no role and, as a result, such models fail to account for the results reported in the present study. Clearly, a full model of lexical processing needs to treat words as more than just a linear concatenation of letters. Not only that, but the model needs to explain how a high frequency bigram can get more confused with a lower frequency one (i.e., *rl* with *lr* in the WALRUS condition) than the other way around (i.e., *mr* with *rm* in the SERMON condition). Such an outcome is quite remarkable and considerably constrains the possibilities for how words are recognized.

Acknowledgment

The research reported in this paper was supported by a grant from the Australian Research Council.

A. Appendix

The following are the items used in Experiments 1 and 2. Each triplet consists of a TL nonword, its one-letter-different control, and its baseword (in uppercase).

<u>NAPKIN condition</u>: nakpin, naktin, NAPKIN; morgtage, morgfage, MORTGAGE; mapgie, mapbie, MAGPIE; tobmoy, tobnoy, TOMBOY; renmant, renbant, REMNANT; semgent, semrent, SEGMENT; hodtog, hodpog, HOTDOG; cobwoy, cobmoy, COWBOY; jawyalk, jawfalk, JAYWALK; sebtack, sebrack, SETBACK; tapdole, tapmole, TADPOLE; bofnire, bofmire, BONFIRE; raibnow, raibhow, RAINBOW; cubpoard, cubtoard, CUPBOARD; domga, domka, DOGMA; gynmast, gynvast, GYMNAST; steafdast, steafbast, STEADFAST; hubsand, hubland, HUSBAND; ruwnay, ruwmay, RUNWAY; vokda, vokma, VODKA

<u>WALRUS condition</u>: warlus, wargus, WALRUS; healdine, healtine, HEADLINE; farbic, fardic, FABRIC; golbin, golpin, GOBLIN; fintess, fincess, FITNESS; berdoom, bergoom, BEDROOM; nulcear, nulgear, NUCLEAR; hooldum, hooltum, HOODLUM; kindap, kingap, KIDNAP; cormade, corvade, COMRADE; tompost, tombost, TOPMOST; paldock, paltock, PADLOCK; oultast, ouldast, OUTLAST; stawlart, stawnart, STALWART; zerba, zerga, ZEBRA; chuntey, chundey, CHUTNEY; sumberge, sumperge, SUBMERGE; hyrdant, hyrfant, HYDRANT; kindey, kingey, KIDNEY; altas, aldas, ATLAS.

<u>SERMON condition</u>: semron, semlon, SERMON; foutnain, foutcain, FOUNTAIN; cidner, cidmer, CINDER; retpile, retfile, REPTILE; boudler, boudner, BOULDER; sybmol, sybnol, SYMBOL; makret, maklet, MARKET; fratnic, fratlic, FRANTIC; racsal, racnal, RASCAL; hadnout, hadmout, HANDOUT; savlage, savhage, SALVAGE; catpive, catbive, CAPTIVE; vetnure, vetmure, VENTURE; stapmede, stapnede, STAMPEDE; padna, padma, PANDA; fatcion, fatsion, FACTION; heatrache, heatwache, HEARTACHE; sevrant, sevlant, SERVANT; capmus, capnus, CAMPUS; hetfy, hetpy, HEFTY.

The following are the items used in Experiment 3. Each triplet consists of a TL nonword, its one-letter-different control, and its baseword (in uppercase). Items marked with an asterisk were also included in Experiments 1 and 2.

<u>NAPKIN condition</u>: *nakpin, nagbin, NAPKIN; *morgtage, morflage, MORTGAGE; *tapdole, tagtole, TADPOLE; *gynmast, gydpast, GYMNAST; laywer, lagmer, LAWYER; rubgy, rupky, RUGBY; zizgag, zicnag, ZIGZAG; diwmit, diyrit, DIMWIT; bamdinton, bawyinton, BADMINTON; wanlut, wamrut, WALNUT; amdiral, anpiral, ADMIRAL; numteg, nunpeg, NUTMEG; hucbap, hudpap, HUBCAP; shranpel, shramdel, SHRAPNEL; stimga, stinfa, STIGMA; absestos, adlestos, ASBESTOS; chinmey, chilrey, CHIMNEY; fafnare, fapvare, FANFARE; goobdye, gootpye, GOODBYE; sihlouette, siwrouette, SILHOUETTE; pocporn, potdorn, POPCORN; fowrard, foylard, FORWARD; fribsee, frigfee, FRISBEE; framgent, franhent, FRAGMENT; anedcote, anetgote, ANECDOTE; pymgy, pynfy, PYGMY; tradpoor, tragboor, TRAPDOOR; cowbeb, coyfeb, COBWEB; hynposis, hymfosis, HYPNOSIS; rendevzous, rendemjous, RENDEZVOUS.

<u>WALRUS condition</u>: *warlus, wascus, WALRUS; *farbic, faldic, FABRIC; *kindap, kistap, KIDNAP; *chuntey, churdey, CHUTNEY; arpon, alcon, APRON; meldey, merney, MEDLEY; manget, marpet, MAGNET; congitive, compitive, COGNITIVE; virbate, vingate, VIBRATE; cultet, curdet, CUTLET; fruiftul, fruimbul, FRUITFUL; golbet, gorset, GOBLET; bulter, burmer, BUTLER; parnter, parster, PARTNER; talbeau, tarpeau, TABLEAU; wintess, wimpess, WITNESS; catherdal, cathental, CATHEDRAL; echinda, echista, ECHIDNA; arcobat, algobat, ACROBAT; mirgant, milcant, MIGRANT; masorny, masoldy, MASONRY; plaftorm, plaskorm, PLATFORM; sharmock, shaldock, SHAMROCK; congac, combac, COGNAC; quardant, qualsant, QUADRANT; chimpunk, chindunk, CHIPMUNK; squardon, squanton, SQUADRON; sumbarine, suptarine, SUBMARINE; pincic, pimbic, PICNIC; mircoscope, mintoscope, MICROSCOPE.

<u>SERMON condition</u>: detnal, decmal, DENTAL; letnil, legril, LENTIL; pitsachio, pinfachio, PISTACHIO; mokney, mogley, MONKEY; vapmire, vatnire, VAMPIRE; bludner, blupser, BLUNDER; cybmal, cypnal, CYMBAL; netcar, nesbar, NECTAR; bewidler, bewitner, BEWILDER; mukset, mupnet, MUSKET; odreal, obseal, ORDEAL; sadnal, sacval, SANDAL; tabmourine, tadnourine, TAMBOURINE; metnion, mecsion, MENTION; shetler, shecner, SHELTER; eanrest, eadlest, EARNEST; petwer, pegder, PEWTER; letcern, lemdern, LECTERN; poproise, podloise, PORPOISE; rubma, rutna, RUMBA; vetrebra, veclebra, VERTEBRA; tetnacle, tepmacle, TENTACLE; capmaign, catnaign, CAMPAIGN; hotsile, hopnile, HOSTILE; gigatnic, gigadric, GIGANTIC; lapmoon, ladsoon, LAMPOON; patnomime, pabsomime, PANTOMIME; tubrine, tugline, TURBINE; podner, pogler, PONDER; vesratile, velnatile, VERSATILE.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, Mass.: MIT Press.

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

Andrews, S., & Scarratt, D. R. (1998). Rule and analogy mechanisms in reading nonwords: Hough Dou peapel rede gnew wirds? *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1052–1086.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*, 12–28.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Berent, I., & Perfetti, C. (1995). A rose is a REEZ: The two cycles model of phonology assembly in reading English. *Psychological Review, 102*, 146–184.

Bruner, J. S., & O'Dowd, D. (1958). A note on the informativeness of parts of words. *Language and Speech, 1*, 98–101.

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

Chen, H.-C., & Vaid, J. (2007). Word frequency modulates the Basic Orthographic Syllabic Structure (BOSS) effect in English polysyllable word recognition. *Language and Cognitive Processes, 21*, 58–82.

Christianson, K., Johnson, R. L., & Rayner, K. (2005). Letter transpositions within and across morphemes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1327–1339.

Davis, C. J. (2006). Orthographic input coding: A review of behavioural evidence and current models. In S. M. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 180–206). Hove, UK: Psychology Press.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117*, 713–758.

Davis, C. J., & Bowers, J. S. (2004). What do letter migration errors reveal about letter position coding in visual word recognition? *Journal of Experimental Psychology: Human Perception and Performance, 30*, 923–941.

Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: Evidence from orthographic similarity effects. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 535–557.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2007). Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho-orthographic decomposition. *Cognition, 105*, 691–703.

Fallows, D. (1981). Experimental evidence for English syllabification and syllable structure. *Journal of Linguistics, 17*, 309–317.

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.

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology, 39A*, 211–251.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers, 35*, 116–124.

Frankish, C., & Barnes, L. (2008). Lexical and sublexical processes in the perception of transposed-letter anagrams. *The Quarterly Journal of Experimental Psychology, 61*, 381–391.

Frankish, C., & Turner, E. (2007). SIHGT and SUNOD: The role of orthography and phonology in the perception of transposed letter anagrams. *Journal of Memory and Language, 56*, 189–211.

Gómez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115*, 577–601.

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). New York: Nova Science.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid falp wrods as a wlohe? *Trends in Cognitive Sciences, 8*, 58–59.

Kay, J., & Marcel, A. J. (1981). One process, not two, in reading aloud: Lexical analogies do the work of nonlexical rules. *Quarterly Journal of Experimental Psychology, 33A*, 397–414.

Lee, C. H., & Taft, M. (2009). Are onsets and codas important in processing letter position? A comparison of TL effects in English and Korean. *Journal of Memory and Language, 60*, 530–542.

Lupker, S. J., Perea, M., & Davis, C. J. (2008). Transposed-letter effects: Consonants, vowels and letter frequency. *Language and Cognitive Processes, 23*, 93–116.

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.

O'Connor, R. E., & Forster, K. I. (1981). Criterion bias and search sequence bias in word recognition. *Memory & Cognition, 9*, 78–92.

Perea, M., & Carreiras, M. (2006). Do transposed-letter similarity effects occur at a syllable level? *Experimental Psychology, 53*, 308–315.

Perea, M., & Carreiras, M. (2008). Do orthotactics and phonology constrain the transposed-letter effect? *Language and Cognitive Processes, 23*, 69–92.

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter confusability effects in masked associative priming. *Memory & Cognition, 31*, 829–841.

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with non-adjacent letter positions. *Journal of Memory and Language, 51*, 231–246.

Perea, M., Rosa, E., & Gómez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception & Psychophysics, 67*, 301–314.

Pulgram, E. (1970). *Syllable, word, nexus, cursus*. The Hague: Mouton.

Rouibah, A., & Taft, M. (2001). The functional role of syllabic structure in French visual word recognition: The BOSS hypothesis revisited. *Memory & Cognition, 29*, 373–381.

Rueckl, J. G., & Rimzhim, A. (2011). On the interaction of letter transpositions and morphemic boundaries. *Language and Cognitive Processes, 26*, 482–508.

Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes, 19*, 333–367.

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention and performance, XII*. London: Lawrence Erlbaum Associates Limited.

Taft, M. (1979). Lexical access via an orthographic code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behavior, 18*, 21–39.

Taft, M. (1987). Morphographic processing. The BOSS re-emerges. In M. Coltheart (Ed.), *Attention and performance, XII*. London: Lawrence Erlbaum Associates Limited.

Taft, M. (1992). The body of the BOSS: Subsyllabic units in the lexical processing of polysyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 1004–1014.

Taft, M. (2001). Processing of orthographic structure by adults of different reading ability. *Language and Speech, 44*, 351–376.

Taft, M. (2002). Orthographic processing of polysyllabic words by native and non-native English speakers. *Brain and Language, 81*, 532–544.

Taft, M., Álvarez, C. J., & Carreiras, M. (2007). Cross-language differences in the use of internal orthographic structure when reading polysyllabic words. *The Mental Lexicon, 2*, 49–63.

Taft, M., & Kougious, P. (2004). The processing of morpheme-like units in monomorphemic words. *Brain and Language, 90*, 9–16.

Taft, M., & Nillsen, C. (2012). *Morphological decomposition and the transposed-letter (TL) position effect*. Language and Cognitive Processes. http://dx.doi.org/10.1080/01690965.2012.679662.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26*, 608–631.

Treiman, R., & Chafetz, J. (1987). Are there onset- and rime-like units in printed words. In M. Coltheart (Ed.), *Attention and performance, XII*. London: Lawrence Erlbaum Associates Limited.

Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language, 29*, 66–85.

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1261–1276.

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.

Whitney, C., & Cornelissen, P. (2005). Letter-position encoding and dyslexia. *Journal of Research in Reading, 28*, 274–301.

Whitney, C., & Cornelissen, P. (2008). SERIOL reading. *Language and Cognitive Processes, 23*, 143–164.