QJEP

What causes the greater perceived similarity of consonant-transposed nonwords?

Teresa Schubert[ID][1], **Sachiko Kinoshita**[ID][1] **and Dennis Norris**[2]

Abstract

Nonwords created by transposing two non-adjacent orthographic consonants (CONDISER) have been reported to produce more priming for their baseword (CONSIDER), and to be classified as a nonword less readily than nonwords created by transposing two orthographic vowels (CINSODER). We investigate the origin of this difference and its relevance for theories of letter position coding. In the unprimed versions of the lexical decision and same–different tasks, a consonant–vowel difference was found in the transposition condition, not when those letters are substituted (Experiment 1). We found that when transpositions involved the disruption of a consonant cluster (OPMITAL), reaction times were slowed compared to when transpositions involved only letters that are separated (CHOLOCATE; Experiment 2). As transpositions more frequently disrupt in consonant clusters than vowel clusters, this introduces a confound in studies investigating consonant and vowel transposition effects. Consistent with the idea that letter order is harder to resolve in clusters, the difference between consonants and vowels was eliminated when transpositions involve singleton consonants or vowels rather than those in clusters (Experiment 3). These results suggest that the precision of position coding does not differ between consonants and vowels, but that consonant–vowel status plays a role in structuring orthographic representations.

An issue that has received much attention in visual word recognition research recently is how letter position is coded within a word or nonword. The empirical phenomenon that stimulated this line of research is the transposed-letter (TL) similarity effect, the finding that a nonword generated by transposing letters in the middle of a word is perceived as highly similar to the baseword. In masked primed lexical decision, a nonword prime constructed by transposing two letters (e.g., jugde) facilitates the recognition of its baseword (JUDGE) more than does a control prime created by substituting the transposed letters with different letters (substituted letter, SL; e.g., junpe). In unprimed lexical decision, TL nonwords are (incorrectly) classified as words more readily, and correctly classified as nonwords more slowly, than nonwords with substituted letters (i.e., the nonword interference effect, e.g., Andrews, 1996; Chambers, 1979; O'Connor & Forster, 1981; Perea & Lupker, 2003; Schoonbaert & Grainger, 2004). These findings challenge the precise coding of letter position assumed in many computational models of visual word recognition[1] (e.g., the interactive-activation model of McClelland & Rumelhart, 1981; the dual-route cascaded model of Coltheart, Rastle, Perry, Ziegler, & Langdon, 2001). Much research has thus been devoted to developing alternative letter position coding schemes. The attempts include various open bigram models (Grainger, Granier, Farioli, van Assche, & van Heuven, 2006; Grainger & van Heuven, 2003; Whitney, 2001; Whitney & Marton, 2013), the spatial coding model (Davis, 2010), the overlap model (Gomez, Ratcliff, & Perea, 2008), and successors to the Bayesian reader model adopting the positional noise

[1]ARC Centre of Excellence in Cognition and its Disorders (CCD), Macquarie University, Sydney, NSW, Australia
[2]MRC Cognition and Brain Sciences Unit, Cambridge, UK

Corresponding author:
Teresa Schubert, ARC Centre of Excellence in Cognition and its Disorders (CCD), Macquarie University, Sydney, NSW 2109, Australia.
Email: teresa.schubert@mq.edu.au

assumption (Norris & Kinoshita, 2012; Norris, Kinoshita, & van Casteren, 2010).

Perea and Lupker (2004; see also Lupker, Perea, & Davis, 2008) have reported a finding that was considered a problem for all of these models: The TL similarity effect was found to be greater for consonants than for vowels. Throughout this paper, the terms "consonant" and "vowel" refer to orthographic consonants and vowels (A, E, I, O, U) rather than phonological ones. The distinction between orthographic and phonological consonant/vowel status has been discussed and empirically demonstrated by a number of authors, as discussed in detail in the General Discussion. Perea, Lupker, and colleagues demonstrated a greater consonant TL similarity effect in a masked primed lexical decision task and in an unprimed lexical decision. In the former task, primes in which (non-adjacent) consonants were transposed (e.g., caniso for CASINO) produced more priming than when they were substituted (e.g., caviro), whereas the vowel-transposed prime condition (e.g., cisano) did not differ significantly from the vowel SL condition (e.g., cesuno). In unprimed lexical decision, TL nonwords were wrongly accepted as words more often, and correctly classified as a nonword more slowly, when consonants were transposed (e.g., caniso) than when vowels were transposed (e.g., cisano). Perea and Lupker (2004) reported these findings with Spanish words, and Lupker et al. (2008) replicated the results in English. Lupker et al. (2008) argued that these results reveal that the position coding of consonants differs from that of vowels, and that this poses a challenge to all extant models of letter position coding because none distinguishes between orthographic consonants and vowels.

A number of recent studies have sought to understand this consonant/vowel effect and clarify its implications for letter position coding. Carreiras, Duñabeitia, and Molinaro (2009) conducted a study of consonant and vowel relative position (subset) primes in the lexical decision paradigm. The subset priming effect refers to the facilitation provided by a prime formed of a subset of letters, in the same relative order, of the target (e.g., apct for APRICOT) over unrelated letters (e.g., egsf for APRICOT). In Spanish, Carreiras et al. (2009) found different event-related potential (ERP) amplitudes (from a 175–250-ms to a 350–450-ms time window) for subset primes formed of consonants (e.g., frl for farol) relative to subset primes formed of vowels (e.g., aeo for acero). On the basis of this result and subsequent follow-up studies revealing an analogous behavioural effect (i.e., facilitation for frl but no difference for aeo compared to an unrelated prime), Carreiras and colleagues proposed the "lexical constraint hypothesis" to explain the advantage of consonant over vowel primes (Carreiras et al., 2009; Duñabeitia & Carreiras, 2011; Vergara-Martínez, Perea, Marín, & Carreiras, 2011). According to this hypothesis, a sequence of consonants is more informative than one of vowels because the former is

consistent with fewer lexical entries—that is, is more lexically constraining. For example, Carreiras et al. (2009) reported that in Spanish the string frl (consonant subset of farol) is consistent with only four words, while aeo (vowel subset of acero) is consistent with 150. They contend that this is true across all languages that have fewer vowels than consonants: Overall, vowels are more frequent than consonants and therefore tend to appear in more words than consonants (Jones & Mewhort, 2004). The relatively small number of lexical entries consistent with a consonant subset prime is assumed to reduce competition, allowing the target to be selected more quickly than when a vowel subset prime is presented.

More recently, New and Nazzi (2014) have extended the lexical constraint hypothesis to primed lexical decision with consonant and vowel substituted-letter primes. They reported that masked primes that preserved the consonants in the target (e.g., duvo–DIVA) facilitated recognition of targets, but primes that preserved vowels (e.g., rifa–DIVA) did not (at 50 ms prime duration; New, Araújo, & Nazzi, 2008; New & Nazzi, 2014). The consonant–vowel difference reported by New and colleagues concerned the effect of substituted-letter primes. In the experiments investigating TL similarity effects, as noted above, SL primes are generally used as the control condition, and the TL similarity effect is defined as the difference between TL and SL conditions. The finding of a consonant/vowel difference in SL conditions raises the possibility that the modulation of the TL similarity effect by consonant–vowel status observed by Lupker and colleagues might reflect a difference in the substituted-letter conditions, rather than in the transposed-letter conditions.

Extant masked primed lexical decision results are consistent with the possibility that difference in the TL similarity effect between consonants and vowels is due to the SL conditions rather than the TL conditions. For example, in Lupker et al. (2008, Experiment 1a), there was a 10-ms difference between the consonant-substituted condition (asiral–ANIMAL; 663 ms) and the vowel-substituted prime condition (anemol–ANIMAL, 653 ms); this difference is numerically equivalent to the difference in the transposed-letter prime conditions (in the opposite direction: consonant-transposed, aminal–ANIMAL: 639 ms; vowel-transposed, anamil–ANIMAL: 650 ms). Thus, the difference between the consonant- and vowel-substituted primes might be explained in terms of the greater lexical constraint provided by the primes that preserved the consonants in the target. For consonant-preserving primes, the few lexical entries consistent with the consonants of the prime are also consistent with the target. For vowel-preserving primes, the numerous lexical entries consistent with the vowels of the prime are also consistent with the target. This produces more competition, which slows recognition of the target when preceded by a vowel-preserving prime compared to a consonant-preserving
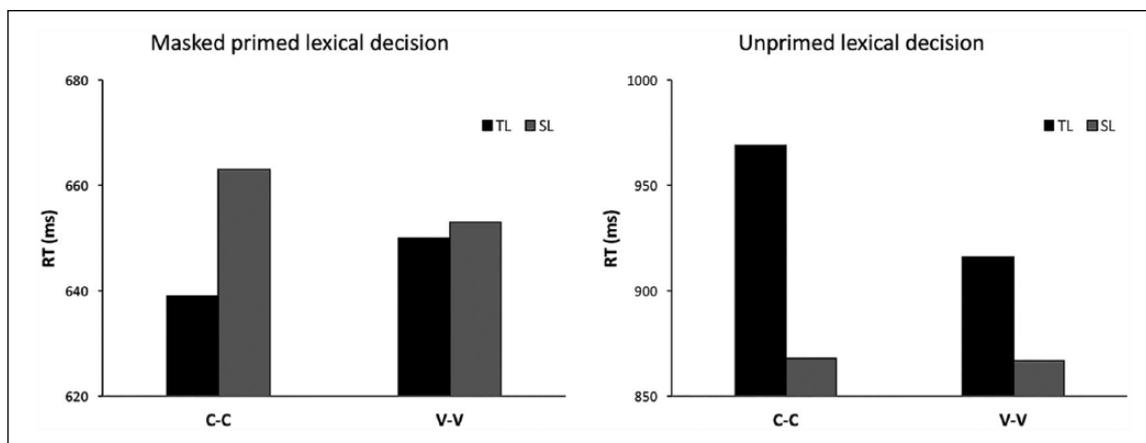
Figure 1. Data from Lupker et al. (2008). Left panel, Experiment 1a (response to word targets preceded by masked primes in the lexical decision task), right panel, Experiment 1b (response to nonword targets in nonword interference task).
Notes: TL = transposed letter, SL = substituted letter; C-C = consonant–consonant, V-V = vowel–vowel; RT = reaction time.

prime. This is an effect of letter identities rather than letter positions.

Perea and Acha (2009) replicated Lupker and colleagues' findings with Spanish items. In addition to a primed lexical decision experiment, they also conducted a primed same–different experiment. In this task, participants are shown the baseword (e.g., casino) as a referent, and are asked to decide whether the target is the same as, or different from, the referent. As in lexical decision, the prime conditions of interest are TL (CANISO, CISANO) and SL (CAVIRO, CESUNO) manipulations for consonants and vowels. Unlike lexical decision, the task does not require lexical access; instead it requires a comparison of orthographic representations of the referent and the target (e.g., Kinoshita & Norris, 2009; Norris & Kinoshita, 2008). In this task, Perea and Acha found no effect of consonant/vowel status on the size of the TL priming effect. The presence of the consonant–vowel effect in lexical decision but not in a same–different task suggests that the origin of the effect is in lexical access, which is obligatory in lexical decision, as Perea and Acha noted. These results can be easily explained as reflecting greater lexical constraint provided by consonant-preserving primes modulating the TL priming effect. The effect is therefore one of letter identity (seen in the substituted-letter conditions) rather than an effect on the coding of letter position (seen in the transposed-letter conditions).

The effect of consonant (C) and vowel (V) transpositions has also been investigated in the unprimed lexical decision paradigm, in which a direct response is required to the transposition and substitution transformed versions of the baseword. Although the C/V status by TL/SL interaction has been observed with both the masked priming lexical decision paradigm and unprimed lexical decision, the pattern of interaction differs between the two tasks.

The difference between these two paradigms is best illustrated with reference to Experiment 1 of Lupker et al.

(2008), which used the same stimuli in both paradigms; their results are graphed in Figure 1. Note that the CC and VV nonwords in their study were generated from the same baseword, so different patterns cannot be attributed to differences in the basewords. The left panel shows the data from nonwords used as masked primes; the right panel shows the reaction times (RTs) to those nonwords when used as targets in unprimed lexical decision. In the unprimed lexical decision task (right panel), there is no difference in the SL conditions, but a large difference between vowels and consonants in the TL condition, with the consonant-transposed nonword (e.g., CANISO) being more difficult to reject as a nonword than the vowel-transposed nonword (CISANO). Crucially, in this task the consonant–vowel difference is observed in the transposition conditions, not the letter substitution conditions. This pattern seems to imply an effect of consonant/vowel status on position coding rather than identity coding, and is not readily explained by the lexical constraint hypothesis (which primarily concerns letter identity). Our aim in this paper is to investigate the basis of the apparent greater similarity of consonant-transposed than vowel-transposed nonwords in the unprimed task, where it suggests an effect of consonant/vowel status on letter position rather than solely letter identity.

We explore this question in the unprimed lexical decision task and the unprimed same–different task; both of these paradigms involve responding directly to the items containing the letter manipulations. In Experiment 1, we show an effect of consonant/vowel status on the perceived similarity of transposed-letter words, which is not due to a difference in the substituted-letter conditions between consonants and vowels. This provides stronger evidence that the effects of letter transposition differ for consonants and vowels. We also report evidence that the difference between consonant and vowel TL conditions cannot be explained in terms of lexical constraints and suggest instead that the difference is due to the preponderance of transpositions

affecting consonant, but not vowel, clusters. In Experiment 2 we directly test the hypothesis that cluster status affects detection of TL manipulations by comparing consonant transposed-letter items that vary in whether they disrupt a cluster. Here we find a robust effect of cluster disruption. Finally, in Experiment 3 we provide additional evidence that the previously observed consonant–vowel difference is attributable to disruption of consonant clusters by demonstrating that the difference disappears when clusters are avoided entirely. We conclude with a discussion of the role of the consonant–vowel status of letters in structuring orthographic representations, and the necessity of representating this information in models of letter position coding.

Experiment 1

In Experiment 1 we compare the unprimed lexical decision task and the unprimed same–different tasks using the stimuli from Lupker et al.'s (2008) Experiment 1. Experiment 1a was an exact replication of their unprimed lexical decision experiment. In Experiment 1b, we used the unprimed same–different match task. In this task, participants are shown the baseword as a referent, and are asked to decide whether the target is the same as, or different from, the referent; the critical targets are TL and SL nonwords. The task does not require finding a lexical entry that matches the target; instead it requires a comparison of orthographic representations of the referent and the target. Accordingly, if the consonant–vowel difference is found only in the lexical decision task, this would suggest that the difference originates in lexical access. To index lexical access, we analysed the data with a covariate of lexical frequency, through linear mixed effects modelling (e.g., Baayen, 2008). We expect that lexical frequency should have an influence in the lexical decision task, but not in the same–different task. If the same pattern of interaction is found in both tasks this would suggest that the consonant–vowel difference does not arise in lexical access.

Method

Participants. Fifty-two volunteer psychology students from Macquarie University participated in Experiment 1 in return for course credit. Twenty-eight performed the lexical decision task (Experiment 1a), and 24 performed the same–different match task (Experiment 1b).

Design. Both Experiments 1a and 1b constituted a 2 (transformation type: TL vs. SL) × 2 (C/V status: consonant vs. vowel) factorial design, with both factors manipulated within subjects. The dependent variables were decision latency (RT) and error rate.

Materials. The stimuli are those used by Lupker et al. (2008, Experiment 1), as listed in the Supplemental Material of

their paper. The critical nonwords were generated from 80 multisyllabic basewords (examples: *CASINO, ACADEMY, MISTAKE*). The basewords were 6–9 letters long (*M* = 7.25 letters), and their mean word frequency (per million) in SUBTLWF (Brysbaert & New, 2009, based on film subtitles) was 14.30 (range = 0.29–101.96). Mean neighbourhood density (N), as defined by Coltheart, Davelaar, Jonasson, and Besner (1977), was 0.34 (range = 0–2), and the mean OLD20 (average orthographic Levenshtein distance to 20 closest substitution, addition, and deletion neighbours, see Yarkoni, Balota, & Yap, 2008) was 2.55 (range = 1.7–3.7). These values are as listed in the English Lexicon Project database (Balota et al., 2002).

The stimuli were 320 nonwords, with four generated from each baseword. All of the nonwords were constructed by changing medial non-adjacent vowels or consonants. In the *vowel-transposed* condition (VVTL), two vowels were transposed (e.g., CISANO). In the *consonant-transposed* condition (CCTL), two consonants were transposed (e.g., CANISO). In the *vowel-substituted* condition (VVSL), the two vowels transposed in the vowel-transposed items were replaced with other vowels (e.g., CESUNO), and in the *consonant-substituted* condition (CCSL), the two consonants were replaced with other consonants (e.g., CAVIRO). (The actual stimuli were not underlined.) The average position of the first transposed/substituted letter was the same for the vowels and consonants (*M* = 3.1 in both cases). Mean OLD20 for the nonwords was 2.86 (range = 1.85 to 4.25), mean N was 0.11 (range: 0 to 3), and mean OrthF (the frequency of neighbours) was 1.01 (range = 0 to 109).

In addition to the nonwords, 80 words were selected to be used as fillers, requiring a *word* response in the lexical decision task (Experiment 1a) and a *same* response in the same–different task (Experiment 1b). They were selected to be similar in lexical characteristics to the 80 basewords.

The 80 basewords were divided into four sets, matched on mean frequency and number of letters. Four list versions were constructed using a Latin square design for the purpose of counterbalancing assignment of 320 nonwords to the four conditions, so that only one of the four nonwords generated from the same baseword (e.g., CANISO, CISANO, CAVIRO, and CESUNO, generated from casino) occurred within a list. There were also 20 practice and initial buffer items selected according to the same criteria as those for the test stimuli. These items were not included in the analysis.

Procedure. Participants were tested in groups of up to three, seated approximately 60 cm in front of a CRT monitor upon which the stimuli were presented. Stimulus presentation and data collection were achieved through DMDX (Forster & Forster, 2003). Stimulus display was synchronized to the screen refresh rate (13.3 ms). The stimuli were presented in Courier New font, in black text on a white background.

In Experiment 1a, each trial started with the presentation of a warning signal consisting of nine # signs for 500 ms. This was then replaced by a target presented in uppercase letters for a maximum of 2000 ms, or until the participant's response. Participants were instructed at the outset of the experiment that on each trial they would be presented with a letter string in uppercase letters, and their task was to decide whether it was a real word, as quickly and accurately as possible.

In Experiment 1b, a reference word was presented (lowercase letters) above the warning signal for 1000 ms. Otherwise, the trial sequence was identical to the lexical decision task. The participant's task was to decide whether the item presented in uppercase letters was the same as or different from the reference word.

Participants were instructed to press a key on a response pad marked "+" for *word* or *same* and a key marked "−" for *nonword* or *different* responses. Participants were given feedback ("Wrong response" message on the screen) when they made an error. Each participant completed 160 test trials (consisting of 80 word and 80 nonword trials in Experiment 1a, or 80 *same* and 80 *different* trials in Experiment 1b), presented in a single block, with a different random order generated for each participant.

Analysis. In this and all subsequent experiments, error rate and RT from the correct trials were analysed using the linear mixed effects model, treating subjects and stimuli as crossed random factors (Baayen, 2008). In the analysis of RTs, we used an inverse transformation (1/RT) to best approximate a normal distribution and meet the distributional assumption of linear mixed effects model, and multiplied by −1000 to maintain the direction of effects and reduce the number of decimal points (i.e., −1000/RT). A cut-off for outliers was determined by inspecting the Q-Q plots of inverse-transformed RT. We used the Lme4 (Version 1.1–5; Bates, Maechler, & Bolker, 2013) package, as implemented in R Version 3.0.3 (R Development Core Team, 2014). Degrees of freedom (estimated using Satterthwaite's approximation)

and *p*-values were estimated using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2013, Version 2.0–11), where possible. If lmerTest failed to compute the *p*-value, we adopted $|t| > 2.0$ as statistically significant, based on the recommendation by Baayen (2008). In line with the recommendation to keep the random effect structure as maximal as consistent with the data (Barr, Levy, Scheepers, & Tily, 2013), the initial model for each analysis included random slopes on participants and items; the final model we report was selected using a backward stepwise model selection procedure. We also entered the interaction between length and the other fixed factors (transformation and C/V status); this was retained where it improved model fit. The factors C/V status ("CV") and transformation type ("TL or SL") were deviation-contrast-coded (−.5, .5) to reflect the factorial design.

Results

Only the *nonword* trials (in Experiment 1a) and the *different* trials (Experiment 1b) were analysed, as the experimental manipulations did not concern the *word* and *same* trials. In Experiment 1a, the outlier cut-off was 350 ms, and two data points (out of a total of 1890) were excluded; in Experiment 1b, the cut-off was 300 ms, and one data point (out of a total of 1530) was excluded. Mean decision latencies and error rates are presented in Table 1.

Experiment 1a (lexical decision task). In the analysis of RTs, the final statistical model has fixed factors of word length, log-frequency of the baseword, previous trial RT (prevRT), transformation type (TL or SL), C/V type (C or V) and their interaction, and subject intercept (28) and word (80) random slope as random factors [invrt ~ prevRT + length + log-SUBTLWF + CV * TLorSL + (stimtype|word) + (1|subject)]. Stimtype is a factor with four levels resulting from a factorial combination of C/V status and TL or SL. The effect of previous trial RT was significant, ß = 0.0002935, SE = 0.0003161, t = 9.29, as was the effect of length,

Table 1. Mean decision latencies and percentage error rates in Experiment 1.

Condition	Letter type							
	Consonant			Vowel			CV	Difference
	Example	RT	%E	Example	RT	%E	RT	%E
Experiment 1a (lexical decision task)								
TL	VADILITY	882	28.9	VILADITY	804	16.3	78	12.6
SL	VABIFITY	728	9.3	VOLEDITY	732	8.0	−4	1.3
TL similarity effect		144	19.6		72	8.3		
Experiment 1b (same–different task)								
TL	VADILITY	636	45.4	VILADITY	600	23.3	36	22.1
SL	VABIFITY	532	6.0	VOLEDITY	526	6.5	6	−0.5
TL similarity effect		104	39.4		74	16.8		

Note: TL = transposed letter; SL = substituted letter; %E = percentage error rate; C = consonant; V = vowel; RT = reaction time, in ms. The mean lexical decision latency to word targets was 690 ms (8.0% errors), and the mean *same* response latency was 529 ms (8.8% errors).

ß=0.04724, *SE*=0.009512, *t*=4.98, with slower RTs being associated with longer words. The effect of baseword frequency was non-significant, ß=−0.01050, *SE*=0.006739, *t*=−1.56. The TL items were responded to significantly more slowly than the SL items, ß=0.1549, *SE*=0.01793, *t*=8.64. Replicating Lupker et al. (2008), the TL manipulation interacted with C/V status, ß=−0.154, *SE*=0.02871, *t*=−4.02. An analysis limited to the TL items showed that VVTL items were responded to significantly faster than the CCTL items, ß=−0.1029, *SE*=0.02314, *t*=−4.45. The effect of baseword frequency was non-significant for the TL items, ß=0.00823, *SE*=0.00999, *t*=−0.82. Length significantly increased RT for the TL items, ß=0.07775, *SE*=0.01386, *t*=5.61.

Error rate was analysed using a mixed effects logit model (Jaeger, 2008), using the same fixed factors as those for the RTs with subjects and item random intercepts, including the interaction with length [errrate ~ logSUB-TLWF + length * CV * TLorSL + (1|word) + (1|subject), family="binomial"], as including the interaction with length improved the model fit: χ^2=23.936, *p*<.001. The effect of length was again significant, ß=0.2064, *SE*=0.0908, *z*=2.273, *p*<.05, with a higher error rate associated with longer words. The effect of baseword frequency was non-significant, ß=−0.0498, *SE*=0.1357, *z*=−0.367, *p*=.71. The TL items were significantly more error prone than the SL items, ß=−3.9883, *SE*=1.1748, *z*=−3.395, *p*<.001. Length interacted significantly with the TL manipulation, ß=0.7071, *SE*=0.161, *z*=4.405, *p*<.001. Across increasing word lengths, the size of the TL effect increased, as can be seen in Figure 2. The presence of this interaction removed the TL × C/V status interaction

that was seen in the RT analysis and when not including the three-way interaction for error rate.

An analysis of only the TL items showed that VVTL items were significantly less error prone than the CCTL items, ß=−0.8799, *SE*=0.1603, *z*=−5.489, *p*<.001. The effect of baseword frequency on TL items was non-significant, ß=−0.005288, *SE*=0.1936, *z*=−0.03, *p*=.98. Length significantly increased the error rate for the TL items, ß=0.5942, *SE*=0.1259, *z*=4.72, *p*<.001.

In sum, Experiment 1a replicated the pattern of data reported by Lupker et al. (2008, Experiment 1b) in RT, showing a greater TL similarity effect for consonant transpositions than for vowel transpositions. In addition, the analysis revealed that there was no effect of baseword frequency, but a robust effect of length (with greater difficulty associated with longer words). In error rate, the interaction between the TL effect and C/V status was modulated by length, such that the TL similarity effect for consonant items increased across length, while the TL similarity effect for vowel items did not.

Experiment 1b (same–different task). In the analysis of RT, the statistical model included the fixed factors of word length, log-frequency of the target word, previous trial RT (prevRT), transformation type (TL or SL), CV type (C or V), and their interaction, with random slopes for words and random intercepts for subjects [invrt ~ prevRT + length + logSUBTLWF + CV * TLorSL + (0 + stimtype|word) + (1|subject)] in the final model. The effect of previous trial RT was significant, ß=0.00039, *SE*=0.000061, *t*=6.50. The effect of length was significant, ß=0.0512, *SE*=0.018, *t*=4.34, with slower RTs
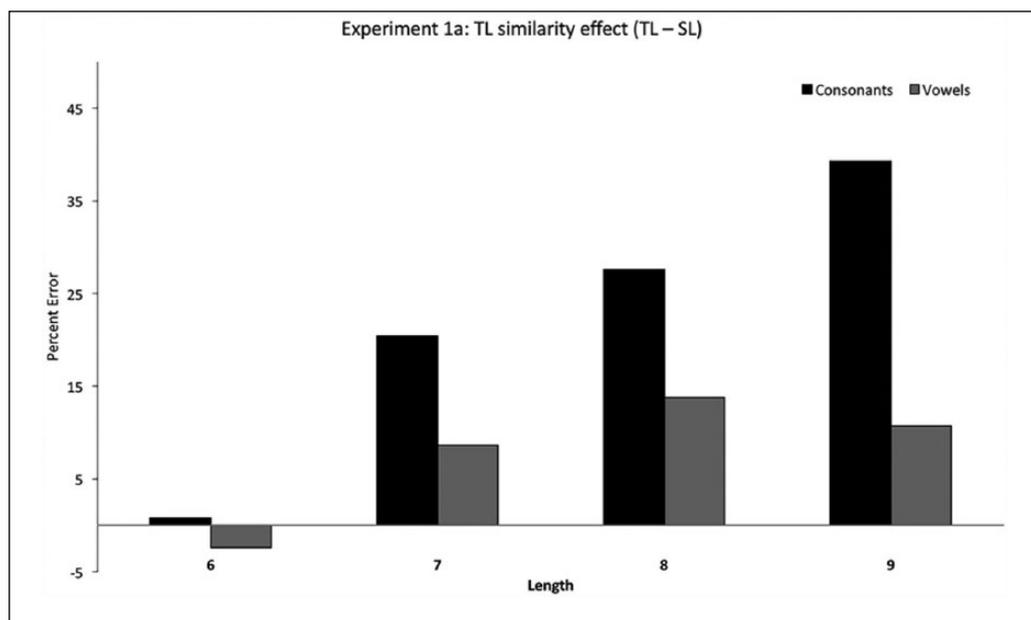

Figure 2. Error rate data from Experiment 1a: Unprimed lexical decision task. The TL similarity effect (TL − SL) is graphed for each length by consonant (CCTL minus CCSL) and vowel (VVTL minus VVSL) items.
Notes: TL = transposed letter, SL = substituted letter; C-C = consonant–consonant, V-V = vowel–vowel.

being associated with longer words. The effect of baseword frequency was non-significant, $\beta = -0.0123$, $SE = 0.00913$, $t = -1.23$. The TL items were responded to significantly more slowly than the SL items, $\beta = 0.0671$, $SE = 0.0174$, $t = 14.03$. Critically, the TL manipulation interacted with C/V status, $\beta = -0.0812$, $SE = 0.0334$, $t = -2.43$. An analysis of only the TL items showed that VVTL items were significantly faster than the CCTL items, $\beta = -0.1107$, $SE = 0.02628$, $t = -4.21$. The effect of baseword frequency was non-significant for the TL items, $\beta = -0.0109$, $SE = 0.0159$, $t = -0.69$. Length significantly increased RT, $\beta = 0.061$, $SE = 0.0202$, $t = 3.04$.

Error rate was analysed using a mixed effects logit model using the same fixed factors as the RT model (except previous trial RT) with random intercepts for subjects and items [errrate ~ length + logSUBTLWF + CV * TLorSL + (1|word) + (1|subject), family = "binomial"]. The effect of length was again significant, $\beta = 0.4191$, $SE = 0.0797$, $z = 5.254$, $p < .0001$, with higher error rate being associated with longer words. The effect of baseword frequency was non-significant, $\beta = 0.01638$, $SE = 0.1241$, $z = 1.32$, $p = .19$. The TL items were significantly more error prone, $\beta = 2.193$, $SE = 0.157$, $z = 13.88$, $p < .0001$. Critically, the TL manipulation interacted with C/V status, $\beta = -1.205$, $SE = 0.310$, $z = -3.89$, $p < .0001$. An analysis of only the TL items showed that VVTL items were significantly less error prone than the CCTL items, $\beta = -1.1535$, $SE = 0.1543$, $z = -7.56$, $p < .0001$. The effect of baseword frequency on TL items was non-significant, $\beta = -0.1975$, $SE = 0.1676$, $z = -1.18$, $p = .24$. Length significantly increased the error rate, $\beta = 0.4325$, $SE = 0.1069$, $z = 4.05$, $p < .001$.

In sum, the pattern of data in the same–different task was very similar to that in the lexical decision task: There was a length effect on both the error rate and RT, there was no effect of word frequency, and it was more difficult to decide that the consonant-transposed nonwords were different from the baseword than the vowel-transposed nonwords.

Discussion

The lexical decision task (Experiment 1a) replicated Lupker et al.'s (2008, Experiment 1b) results: Nonwords generated by transposing consonants (e.g., *CANISO*) were classified as nonwords more slowly than nonwords generated by transposing vowels (e.g., *CISANO*). Using the same items in the same–different task (Experiment 1b), which does not require lexical access, we found the same pattern. This was unexpected based on the lexical constraint theory of Carreiras et al. (2009), which predicts a C/V effect only in the lexical decision task.

The linear mixed effects analysis that we used offers some insight: It revealed that the response to the TL nonword targets (for both RT and error rate) was insensitive to baseword frequency[2] (range 0.29–101 per million in SUBTLEX frequency), but sensitive to word length (range

6–9 letters), in both the lexical decision task and the same–different task. Furthermore, the *nonword* responses were very slow. The TL nonwords produced RTs over 800 ms; in Lupker et al.'s original experiment (Lupker et al., 2008, Experiment 1b), RTs were well over 900 ms.[3] The lack of a word frequency effect, combined with a word length effect and very long RTs on the TL nonwords in the lexical decision task suggests that the *nonword* responses in this task were not a pure reflection of the ability of TL nonwords to "activate . . . the lexical representation of their base words" as suggested by Perea and Lupker (2004, p. 236). Instead, we suggest that responding *nonword* in the lexical decision task involved a detailed orthographic comparison, just as in the *different* response in the same–different task. In the same–different task the comparison is between target and referent, while in lexical decision it is between the nonword target and its baseword. The involvement of detailed orthographic comparison in the unprimed lexical decision task is perhaps not surprising given that all of the TL nonword targets were generated from long polysyllabic words with few neighbours (e.g., CONSIDER, MILITARY). We suggest that such a comparison would result in long RTs and mask any baseword frequency effect.

The question then becomes: Why does it take longer to find a difference between a consonant-transposed nonword and the baseword than between a vowel-transposed nonword and the baseword? We explore this question in the following two experiments, utilizing the same–different task as it unambiguously involves the type of detailed orthographic comparison that we are interested in.

Experiment 2

One clue to the differing effect of consonant and vowel transpositions comes from the error analysis of Experiment 1a. There we found a significant three-way interaction between length, transformation (TL or SL), and C/V status, with the size of the TL similarity effect increasing with length for consonant transpositions but not for vowel transpositions (Figure 2). Curious as to what might be driving the increase in the consonant items, we inspected the items across length and found that while almost all of the six-letter items (e.g., MEMORY, ANIMAL) contained purely consonant and vowel singletons (i.e., CVCVCV or VCVCVC), longer words were more likely to contain clusters, particularly consonant clusters (e.g., OPTIMAL, DENSITY, CARDINAL). Could the increase in consonant clusters in the longer items (not paralleled by increased numbers of vowel clusters) explain the TL similarity effect, which was particularly evident for the long consonant items?

Examination of the stimuli that produced high error rates across Experiments 1a and 1b suggests a pattern consistent with this. In Experiment 1a, examples of consonant-transposed nonwords with high error rates are

SENMITENT (71.4%), CONDISER (71.4%), and OPMITAL (57.1%; see Supplemental Material A for the full list); the transpositions in these items disrupt consonant clusters. Likewise in Experiment 1b, many of the same items produced high error rates (e.g., SENMITENT, 83.3%). The mean error rate for the disrupted-cluster items was higher for the consonant-cluster disrupted items (32.7%) than for those in which singletons were transposed (28.1%) in both experiments (1a: 32.7% vs. 28.1%; 1b: 47.6% vs. 44.9%).[4] In contrast, transposed vowels rarely disrupted clusters (only 3 TL manipulations affected a vowel cluster). The discrepancy between small numbers of vowel cluster-disrupting and large numbers of consonant cluster-disrupted items also appears to be the case with the items of Perea and Lupker (2004).[5] For example, in their Experiment 4, 14 out of 40 consonant transpositions disrupted a consonant cluster (e.g., ESREPANZA from esperanza; GUSBATA from gustaba; JUSCITIA from justicia), but no vowel transpositions disrupted a vowel cluster.

The difficulty of locating a consonant or a vowel embedded within a same-category cluster has been reported previously. For example, using a letter search paradigm, Brand, Giroux, Puijalon, and Rey (2007, Experiment 2) reported that a consonant was harder to detect in a multi-letter syllable onset of a French word (e.g., L in TABLIER) than in a single-letter syllable onset matched in position (e.g., L in ECOLIER). The same difficulty was found with a vowel occurring within a multi-letter grapheme (e.g., U in BOULE) relative to a single-vowel context (e.g., U in BRUNE; Experiment 2). Could the transposition of letters from a consonant cluster explain the apparent greater perceived similarity of consonant-transposed nonwords to the basewords?

To test this hypothesis, we conducted an experiment to directly compare items containing a transposition that disrupts a consonant cluster to those containing a transposition of singleton consonants. Though we expect that orthographic processing of vowel clusters also differs from processing of vowel singletons, only consonant items were used, as vowel clusters are rarer and often comprise graphemes (e.g., TAUT). As the primary comparison is between the cluster and singleton items, we simplified the design and length of the experiment by limiting the critical stimuli to TL and not including an SL manipulation.

Method

Participants. Twenty-nine psychology students from Macquarie University participated in Experiment 2 in return for course credit.

Design. Experiment 2 used the same–different match task, where all of the different items involved consonant transpositions that spanned an intervening vowel. The critical manipulation concerned whether the letter transposition affected a consonant cluster (e.g., ALHOCOL) or only singletons (e.g., LUTANIC). The dependent variables were decision latency and error rate.

Materials. The critical stimuli were 96 nonwords generated by transposing consonants in 7- and 8-letter-long multisyllabic words. Half of the words contained consonant clusters in internal positions (e.g., alcohol, artisan), and the other half did not (e.g., lunatic, lateral). They were matched on length (mean 7.54 letters), mean frequency (per million SUBTLWF, 5.18, range = 0.06–33.2), N (*M* = 0.19, range = 0–2), and OLD20 (*M* = 2.80, range = 2.05–3.95).

To construct the critical nonwords, non-adjacent consonants (spanning a single vowel) in word-internal positions were transposed—for example, ALHOCOL, LUTANIC. The position of the transposed letters was matched between the cluster and singleton words (the first transposed letter always occurred in either the third or fourth position, mean 3.1). Mean OLD20 for the nonwords was 3.06 (range = 2.2 to 3.85), mean N was 0.05 (range = 0 to 2), and mean OrthF was 0.28 (range = 0 to 12.43).

The 96 words were divided into two sets, matched on length. Two list versions were constructed for the purpose of counterbalancing assignment of sets to the *same* and *different* conditions, so that within a list, each target word occurred only once and, across the two lists, appeared once intact as a word requiring the *same* response, and once as a transposed-letter nonword, requiring a *different* response. In addition, there were 20 practice and initial buffer items, selected according to the same criteria as those for the test stimuli. These items were not included in the analysis.

Procedure. The apparatus and general procedure were identical to those of Experiment 1b. Each participant completed 96 test trials (consisting of 48 *same* and 48 *different* trials), presented in one block, with a different random order generated for each participant. The stimuli were presented in Courier New font, in white text on a black background.

Results

We first report the analysis of the *different* responses as in the previous experiments. The general method of analysis was identical to that of Experiment 1b. All RTs were greater than 350 ms and therefore surpassed the outlier cutoff (a total of 1094 data points). Mean decision latencies and error rates are presented in Table 2. (Supplemental Material B lists the critical stimuli and their error rates.)

In the analysis of RTs, the statistical model included as fixed factors previous trial RT, baseword log-frequency, length, and item type (cluster or singleton), with subject (29) and word (96) intercept as random factors [invrt ~ prevRT+ logSUBTLWF+length+itemtype + (1|word) + (1|subject)]. As in the previous same–different experiment, previous trial RT was highly significant, $\beta = 0.000178$, *SE* = 0.00040, $t = 4.40$, and the effect of baseword frequency was

Table 2. Mean decision latencies and percentage error rate in Experiment 2.

Condition	Item type							
	Cluster			Singleton			Difference	
	Example	RT	%E	Example	RT	%E	RT	%E
TL	ALHOCOL	837	21.4	LUTANIC	806	21.4	31	0
Same	ALCOHOL	770	8.2	LUNATIC	725	6.4	45	1.8

Note: TL = transposed letter; %E = percentage error rate; RT = reaction time, in ms. Same–different match task.

non-significant, ß = −0.00376, *SE* = 0.0101, *t* = −0.31. Although the responses were numerically slower in eight-letter words than in seven-letter words (798 ms vs. 837 ms), the factor length was non-significant, ß = 0.03627, *SE* = 0.02585, *t* = 1.40, probably due to the limited range. Critically, the TL items were responded to significantly faster when the transposition occurred in a singleton context than when it occurred in a cluster (main effect of item-type), ß = −0.05832, *SE* = 0.02652, *t* = −2.20. Error rate was analysed using a mixed effects logit model, using the same fixed factors as those for the RTs [errrate ~ logSUB-TLWF + length + itemtype + (1|word) + (1|subject)] (family = "binomial"). None of the fixed factors reached significance, all |*t*| < 1.304.

As can be seen in Table 2, the *same* responses showed the same critical effect of cluster type as the *different* responses. This is confirmed in the absence of an interaction between the item type and the response type factors (ß = .0115, *SE* = 2.737e-02, *t* = 0.420) in an analysis that included both the *same* and *different* responses [invrt ~ prevRT + logSUB-TLWF + length + resptype * itemtype + (1|word) + (1|subject)]. The itemtype effect was significant, indicating that cluster items were responded to more slowly than singleton items: ß = −.06108, *SE* = 2.017e-02, *t* = −3.028. For error rate also, the resptype by itemtype interaction was non-significant: ß = 0.2848, *SE* = 0.2525, *z* = 1.128.

Discussion

The results of Experiment 2 provide evidence that transpositions that disrupt a consonant cluster are harder to detect. Cluster TL items were responded to 31 ms slower than singleton TL items, suggesting difficulty in locating the transposed letter when it was in a consonant cluster; this was a direct test of whether cluster disruption affects processing. Because the *different* trials manipulate transpositions in clusters and singletons we can be sure that the difference specifically reflects the CV structure of the words.

Words in the *same* trials have the same CV properties [cluster (ALCOHOL) and singleton (LUNATIC) base-word items], and thus the difficulty in processing clusters was also expected (and obtained) in this condition. This is due to the structure of the experiment: Because there were no SL items present, and because it is clearly impossible to

know ahead of time whether a given trial is *same* or *different*, it is not sufficient to simply identify the letters present in the stimulus. Instead, on every trial one must code both letter identities and positions accurately enough to distinguish between ALHOCOL and ALCOHOL; we contend that this position coding process is sensitive to CV structure, and slowed by the presence of clusters. Therefore, the difference between ALCOHOL and LUNATIC in the *same* trials is interpreted in the same way as the difference between ALHOCOL and LUTANIC in the *different* trials: difficulty in localizing a consonant within a consonant cluster.

When TL items are not controlled for the disruption of clusters, the presence of consonant cluster transpositions (but not vowel cluster tranpositions) may explain the observed difference between consonant and vowel transpositions. However, when TL items affect only singleton letters, we expect no difference to emerge between consonant and vowel transpositions. This prediction was tested in Experiment 3.

Experiment 3

If the difficulty in locating a transposed consonant from within a consonant cluster is the cause of the apparent greater perceived similarity for consonant-transposed nonwords, then the difference between consonants and vowels should disappear when all letter transpositions occur in singleton contexts. We selected a new set of words containing a medial CVCV or VCVC segment (e.g., chocolate), and non-adjacent consonants or vowels were transposed within the segment (e.g., CHOLOCATE, CHOCALOTE).

Method

Participants. Twenty psychology students from Macquarie University participated in Experiment 3 in return for course credit.

Design. Experiment 3 constituted a 2 (transformation type: TL vs. SL) × 2 (C/V status: consonant vs. vowel) factorial design, with both factors manipulated within subjects. The dependent variables were decision latency and error rate.

Table 3. Mean decision latencies and percentage error rates in Experiment 3.

Condition	Letter type							
	Consonant			Vowel			CV Difference	
	Example	RT	%E	Example	RT	%E	RT	%E
TL	CHOLOCATE	920	43.2	CHOCALOTE	901	36.8	19	6.4
SL	CHOSORATE	638	6.9	CHOCULITE	630	5.6	8	1.3
TL similarity effect		282	36.3		271	31.2		

Note: TL = transposed letter; SL = substituted letter; %E = percentage error rate; C = consonant; V = vowel; RT = reaction time, in ms. The mean *same* response latency was 684 ms (10.3% errors).

Materials. The critical stimuli were 320 nonwords generated from 80 multisyllabic words. The words were all nine letters long, and their mean frequency (per million, SUBTLWF) was 4.49 (range = 0.02–140.67). Mean N was 0.21 (range = 0–1), and OLD20 was 3.0 (range = 2.25–3.90). These stimuli were long (9 letters) to avoid making a change to the first syllable.[6] The 80 words used as fillers (also 9 letters long) requiring the *same* response were selected as in Experiment 1b.

The construction of four types of stimuli—namely, VVTL, CCTL, VVSL, and CCSL—was similar to that of Experiment 1. The transpositions involved non-adjacent letters in internal positions. Position of transposed letters and the distance between them was matched between the consonant- and vowel-transpositions (mean 4.7th position for the first transposed letter). The transposed consonant/vowel never occurred within the first syllable, or within a cluster of consonants/vowels, e.g., SIGNA<u>R</u>U<u>T</u>E, not SIG<u>T</u>A<u>N</u>URE.[7] Mean OLD20 for the nonwords was 3.57 (range = 2.65–4.85), mean N was 0.02 (range = 0–1), and mean OrthF was 0.08 (range = 0–14.22).

The 80 items were divided into four sets, matched on mean frequency of the baseword. Four list versions were constructed for the purpose of counterbalancing assignment of sets to the four conditions, so that within a list, each target word occurred only once and, across the four lists, appeared in all four experimental conditions. In addition, there were 20 practice and initial buffer items, selected according to the same criteria as those for the test stimuli. These items were not included in the analysis.

Procedure. The apparatus and general procedure were identical to those of Experiment 1b. Each participant completed 160 test trials (consisting of 80 *same* and 80 *different* trials), presented in a single block, with a different random order generated for each participant. The stimuli were presented in Courier New font, in black text on a white background.

Results

The general method of analysis was identical to that in Experiment 1b. Six data point outliers (from a total of 1228) with RTs faster than 350 ms were excluded from the analysis. Mean decision latencies and error rates are presented in Table 3. (Supplemental Material C lists the critical stimuli and their error rate.)

In the analysis of RTs, the statistical model included as fixed factors previous trial RT, baseword log-frequency, transformation type (TL or SL), C/V type (C or V), and their interaction, and subject (20) and word (80) slopes as random factors [invrt ~ prevRT+logSUBTLWF+CV * TLorSL+(0+stimtype|word)+(0+stimtype|subject)]. The effect of previous trial RT was significant, ß=0.0002074, $SE=0.0003871$, $t=5.36$. As in the previous experiments, the effect of baseword frequency was non-significant, ß=−0.00903, $SE=0.01718$, $t=−0.53$. The TL items were responded to significantly more slowly than the SL items, ß=0.3725, $SE=0.04252$, $t=8.76$.

Direct comparison of VVTL and CCTL items was non-significant, ß=−0.005128, $SE=0.0364$, $t=0.141$. Furthermore, the TL manipulation did not interact with C/V status, ß=0.0213, $SE=0.0508$, $t=0.419$. To quantify the amount of evidence for the null interaction, we calculated the Bayes factor (using the BayesFactor package v 0.9.7, "compare" function with default JZS prior in R, Morey & Rouder, 2013). We compared the model with the TL/SL by C/V status interaction (Model 1) with a model that did not include the interaction (Model 2). A Bayes factor of 1 indicates equal evidence for the hypotheses (they are equally plausible), less than 1 indicates more evidence for Model 1, and greater than 1 indicates more evidence for Model 2, with a Bayes factor of 3 considered "some evidence" according to Jeffreys (1961) classification. The Bayes factor for our comparison was 5. Thus, this analysis indicated that there was reasonable evidence for the null interaction between consonant/vowel status and TL/SL.

Error rate was analysed using a mixed effects logit model, using the same statistical model as that for the RTs excluding the prevRT factor [errrate ~ logSUBTLWF+CV * TLorSL+(1|word)+(1|subject), family ="binomial"]. The effect of baseword frequency was non-significant, ß=−0.20791, $SE=0.1169$, $z=−1.778$, $p=.075$. The TL items were significantly more error prone, ß=2.558, $SE=0.1827$, $z=14.002$, $p<.0001$. As with RT, the difference between VVTL and CCTL items was non-significant,

ß = −0.2721, *SE* = 0.1741, *z* = −1.563, *p* = .12. Consistent with the RT data, the TL manipulation did not interact with C/V status, ß = 0.08194, *SE* = 0.3482, *z* = 0.235, *p* = .81. As it is currently not possible to calculate the Bayes factor with logit models, we simply compared models that differed only in the inclusion of the interaction to see whether the inclusion of the interaction improved the data fit. It did not, χ^2 = 0.0577, *p* = .82.

Discussion

In Experiment 3, all transposed consonants and vowels were singletons, and the consonant–vowel difference was eliminated. Whilst a highly robust TL similarity effect was observed both with latency and with error rate, this effect was not modulated by consonant–vowel status. Direct comparison between the consonant-transposition and vowel-transposition conditions showed no difference in either error rate or latency. Combined with the results of Experiment 2, we take these results to suggest that the difference observed between consonant- and vowel-transpositions in Experiment 1 were due to the difficulty of locating a transposed consonant within a multi-consonant cluster.

General discussion

We set out to investigate the origin of the consonant–vowel difference in the transposed-letter (TL) similarity effect first reported by Perea and Lupker (2004). Experiment 1 used nonwords generated by transposing nonadjacent consonants and vowels in polysyllabic English words (primes from Lupker et al., 2008, Experiment 1) in an unprimed lexical decision task and same–different task. We replicated the greater TL similarity effect with consonant transpositions in lexical decision reported by Lupker et al. (2008) and extended the finding to the same–different task. Unlike in masked priming experiments (e.g., Lupker et al., 2008; Perea & Acha, 2009), which appear to show a consonant–vowel difference in the substituted-letter conditions, here the consonant–vowel difference was found only in the transposed-letter conditions. This result is a clearer demonstration of an effect of C/V status on letter position coding than those previously reported. The fact that "nonword" responses to the TL nonwords were insensitive to baseword frequency in this experiment suggests that the lexical constraint hypothesis may not explain this C/V effect. Instead, we propose that the TL similarity effect in both tasks reflects the difficulty in distinguishing between the orthographic representation of the baseword (retrieved from the lexicon in the case of the lexical decision task, and encoded from the presented referent in the same–different task) and the nonword target. We used the unprimed same–different task to investigate the source of the greater perceived similarity of consonant-transposed words in two further experiments.

Based on examination of the materials from Experiment 1, we hypothesized that the presence of consonant clusters might present a challenge to locating a letter and evaluating a transposed item compared to its referent. In Experiment 2 we compared the effect of transpositions in singleton and cluster items, and found that participants were slower to reject TL items that contained a change to a consonant cluster than when the change occurred to consonant singletons. Finally, in Experiment 3 we used items where the transposed consonants (and vowels) were always singletons (e.g., CHOLOCATE, CHOCALOTE) rather than from within a consonant cluster (e.g., CONDISER), and the difference between consonant and vowel transpositions was eliminated. This result indicated that the greater effective similarity of consonant-transposed words in Experiment 1 was not due to an intrinsic property of consonant or vowel letters, but due to the context in which they appear.

Locating a consonant is more difficult when it occurs within a cluster than when it occurs as a singleton. The greater TL similarity effect observed with consonants in previous studies using unprimed lexical decision with English and Spanish polysyllabic words is also consistent with this observation, as in these studies the consonant-transpositions often occurred in a consonant cluster whereas vowel transpositions rarely occurred in a vowel cluster (refer to the list of stimuli in Supplemental Material A and Perea & Lupker, 2004). The conclusion of a cluster effect on letter position coding would be further supported by similar manipulations of singleton/cluster structure and consonant/vowel status—for example, an omnibus experiment testing the full factorial combinations of the orthographic properties that are thought to be relevant. One difficulty that might arise with this approach is stimulus selection, as the confound between consonant/vowel status and clusters seems to arise quite naturally in a number of alphabetic languages. We look forward to future work that addresses this issue. This confound seems to arise quite often in polysyllabic words because consonant clusters are generally more common than vowel clusters in these words (particularly in languages where vowel digraphs are rare, e.g., Spanish).

The disappearance of the consonant–vowel difference when the transposition involved a singleton consonant/vowel suggests that the precision of position coding of a consonant letter is no different from a vowel letter. This is consistent with the conclusion drawn by Perea and Acha (2009) based on their masked priming results, but their view that "letter position coding occurs before the consonant/vowel distinction begins to matter" (p. 136, and similar statement on p. 135) would have trouble explaining why a consonant–vowel difference was observed in the unprimed same–different task (Experiment 1b). What the present results indicate is that the difference between the baseword and the nonword generated by a transposing non-adjacent letter (consonant or a vowel) is harder to

perceive when one of the transposed letters is embedded in a cluster. This result implies that the consonant–vowel category status of letters is represented in pre-lexical orthographic representations.

The idea that consonant–vowel status is used to structure an orthographic representation is not new and has been suggested across a variety of tasks. Acha and Perea (2010) conducted a letter search task (in Spanish) with words and pseudowords, which revealed that the ease of detecting consonants and vowels depends on their position and is linked to the positional distribution of consonants and vowels in the language. These results are consistent with the use of a CV skeleton in pre-lexical representations, biasing the search for consonant and vowel letters to particular positions where they are frequently found. As noted above, Brand et al. (2007) found that it is harder to detect a consonant when it occurred in the second position of a complex onset (e.g., L in TABLIER) than when it occurred as a simple onset (e.g., L in ECOLIER). Chetail and Content (2013) showed that this effect was specifically orthographic: it was easier to detect a consonant within a syllable onset cluster if an E intervenes, even when phonologically silent (e.g., detecting R was easier in BANDERILLE /bãdʀij/ than in ESPADRILLE /ɛspadʀij/; note that the phonological syllable containing the R is identical). Chetail, Drabs, and Content (2014) conducted an unprimed same–different task in which two letters of the target were transposed: two consonants, two vowels, or one consonant with one vowel. Only the third condition involves a disruption to the CV-structure of the target, and reaction times in this condition were faster than those in the other two, suggesting facilitation in detecting a letter change when it entails a change to the CV-structure. Furthermore, when the transposition disrupted a consonant cluster, participants were slower to detect the change than when it did not, which is consistent with the results of our Experiment 2. Use of the same–different task in Chetail's work as well as our own provides further support for CV structure having a pre-lexical locus. Clearly, from across a variety of tasks there is evidence that the consonant–vowel status of the letters structures orthographic representations, particularly when the input is a long polysyllabic word like the stimuli used here.

Our findings also relate to the notion of an orthographic CV tier discussed in the context of spelling (e.g., Buchwald & Rapp, 2006; Caramazza & Miceli, 1990; McCloskey, Badecker, Goodman-Schulman, & Aliminosa, 1994; Miceli, Capasso, Benvegnù, & Caramazza, 2004). For example, Buchwald and Rapp (2006) studied the spelling errors produced by patients with deficits to the graphemic buffer, or orthographic working memory. In spelling to dictation, these individuals' errors often preserved the consonant–vowel status of the word—for example, CHAIN is likely to be misspelt as CHAON or STAIN, but not CHALN. This result suggests a multidimensional/hierarchical representation in orthographic working memory

comprising both letter identities and orthographic consonant/vowel status, with the possibility of a deficit affecting just the former. Though these particular studies considered spelling—that is, orthographic output—some authors suggest that the graphemic buffer is also employed during reading as the temporary store of letter identities and their positions (e.g., Forde & Humphreys, 2005; Schubert & McCloskey, 2015; Tainturier & Rapp, 2003). As such, the graphemic buffer is a plausible candidate to maintain orthographic information during comparison between the TL nonword target and the baseword, as in our experiments. Our finding that consonant-transpositions are harder to detect when they involve a consonant cluster is consistent with a structured orthographic representation at this level. Next, we review whether any existing models of visual word recognition are sensitive to orthographic CV structure and could account for our findings.

Implications for models of orthographic processing

It has been suggested (Lupker et al., 2008; Perea & Lupker, 2004) that the consonant–vowel difference in the TL similarity effect presents a challenge for recent theories of orthographic processing because they do not distinguish between consonants and vowels. Indeed, none of these proposals—the various open bigram models (Grainger et al., 2006; Grainger & van Heuven, 2003; Whitney & Marton, 2013), the spatial coding model (Davis, 2010), the overlap model (Gomez et al., 2008), the graded both-edges theory (Fischer-Baum, Charny, & McCloskey, 2011; McCloskey, Fischer-Baum, & Schubert, 2013), and the noisy channel model (Norris & Kinoshita, 2012)—make a distinction between consonants and vowels at the letter level. Insofar as no consonant–vowel difference was observed when the transposition involved singleton consonants/vowels (Experiment 3), this may be taken to suggest that no modification is required to this assumption. However, our results also imply that identifying a consonant is more difficult in clusters of consonants, and thus the presence of clusters (i.e., adjacent same-category letters) must be encoded.

This notion seems to pose the greatest challenge for open-bigram models. Open bigrams are ordered letter pairs that may be contiguous or non-contiguous, and open-bigram models posit that a word is represented as an unordered set of open bigrams (e.g., WORD is represented as [RD], [OD], [OR], [WR], [WO]). In open-bigram models, transposed-letter effects are explained by the fact that transposing letters will alter fewer bigrams than will substituting letters. However, open bigrams cannot explain the cluster effect that we observed. For example, the cluster-pairs INHIBIT and INBIHIT share 18 of their 21 open bigrams, as do the non-cluster pairs HABITAT and HATIBAT. In both cases they differ in the bigram of the transposed letters (HB–BH and BT–TB) and two other

bigrams each containing a vowel and one of the transposed consonants (IB–BI, HI–IH and IB–BI, IT–TI). That is, the bigrams that differ have the same CV structure regardless of whether the transposition involves letters in a cluster. Some open bigram theorists posit that in addition to open bigrams, a "fine-grained" orthographic code is also employed to code letter order precisely (e.g., Grainger & Hannagan, 2014). This fine-grained route might have the capacity to represent cluster status; however, it is incompatible with TL similarity effects in general and therefore cannot conceivably underlie a TL × Cluster interaction.

The notion of a CV structure can be reconciled more readily with models that employ abstract letter units, rather than bigrams. One such model is the spatial coding model, in which letter position is coded as an activation gradient with the letter in the first position assigned the highest (or lowest) level of activation, the letter in the second position the next highest (lowest), and so on (Davis, 2010). Other models assume that an orthographic representation consists of letter representations whose locations within a word are noisy (e.g., overlap model, noisy channel model, graded both edges). Functionally structured orthographic representations based on the consonant–vowel status of the letters are not a priori incompatible with these models because they use letters as the unit of orthographic representation, the same grain size as that at which consonant–vowel status is relevant.

More recently, Chetail and colleagues (e.g., Chetail et al., 2014; Chetail, Treiman, & Content, 2015) have posited that consonant/vowel status of letters is directly utilized in word recognition. They propose that information about abstract letter identities and their positions is subsequently represented by vowel-centred units. For example, the word GALA would be represented initially by its constituent letters ([G], [A], [L], [A]) and then by two units corresponding to the vowels ([GA], [LA]). [These units need not correspond to the phonological syllables: They demonstrated that words like CHAOS (two phonological syllables) have one vowel-centred unit.] Lexical access then occurs on the basis of the vowel-centred units. Chetail and colleagues' CV pattern theory is the first theory of visual word recognition to explicitly include orthographic units that depend on a letter's C/V status, but at this stage it includes only minimal details about representation of letter position and thus cannot account for TL similarity effects. As presently specified then, no existing letter-based models can account for the TL by cluster interaction that we report here.

We do not take a strong position on whether the specific representations (e.g., CV skeleton, orthographically defined CV syllables) proposed in these studies are the perceptual units driving lexical access. However, they are plausible proposals for how C/V information may structure orthographic representations. Future theories should aim to account for effects of both C/V status and letter transposition. Whether the various C/V effects reported in the literature are better understood as arising due to a level of vowel-centred units, by representation of the C/V status of individual letters within the graphemic buffer, or by another theory yet to be proposed, remains to be determined.

Conclusions

The present study yields two novel insights: First, we pointed out that the greater TL similarity effect found with consonants than vowels in masked priming studies may reflect the difference between substituted-letter conditions rather than transposed-letter conditions. Second, we demonstrated that the apparent greater similarity of nonwords generated by transposing consonants is due to the difficulty of locating a transposed consonant when it is a constituent of a consonant cluster. Disrupting a consonant cluster was found to slow processing of TL items, and when a letter transposition instead involved a singleton consonant or vowel, no difference was observed between consonant-transposed and vowel-transposed nonwords. These findings indicate that there is no inherent difference in the precision of position coding for individual consonants and vowels, though consonants within a letter string provide more lexical constraint than vowels. The finding that letter transpositions are harder to perceive when the transposed consonant occurs in a cluster of consonants suggests that orthographic representations (particularly of polysyllabic words) are not just a linear string of letters, but appear to be structured according to orthographic consonant–vowel status. We believe the findings presented here pave a way forward for current models of orthographic processing to address the role of consonant–vowel status in coding letter position.

ORCID

Teresa Schubert http://orcid.org/0000-0002-1627-3893
Sachiko Kinoshita http://orcid.org/0000-0002-5288-2485

Notes

1. It should be noted, however, that the use of precise position-specific codes in these models was as much a matter of implementational convenience rather than a consequence of a strong theoretical commitment.
2. The absence of a frequency effect on the nonword responses is not due to a weak manipulation of frequency; in a version of Experiment 1a run as a masked priming lexical decision task (using the nonwords as primes for their basewords) we found a robust frequency effect ($p < .0001$).
3. Furthermore, the RTs in the experiment described in Footnote 2, which used nonwords generated by changing two or more letters of existing words (e.g., ORPHISECT from architect; BERADE from female) generated mean RTs of 539 ms, over 200 ms faster.
4. Note that these comparisons involve different items, and are not equated on other factors such as length and whether the

letter transposition changes pronunciation (e.g., LOCIGAL, but not CONDISER, changes the pronunciation of "C"). These other factors are also likely to contribute to the ease of detecting the difference from the baseword.

5. In the Appendix of Perea and Lupker (2004), the stimuli used in Experiment 3 (and 4) are labelled incorrectly (consonant-transposition should be vowel-transposition and vice versa).

6. The stimuli of Lupker et al. (2008, Experiment 1) used in our Experiment 1 involved a change in the first syllable in 37 vowel transpositions, but none of the consonant transpositions (see Supplemental Material A). Whatever the reason (e.g., greater attention, reduced crowding), altering the first syllable may make the change more salient. This factor alone cannot account for the results reported by Perea and Lupker (2004) because they also reported an interaction between C/V status and SL/TL in one experiment with no changes to the first syllable (Experiment 4), but we thought it best to avoid the potential confound.

7. Due to experimenter error, a few of the items (discovery, endurable, incidence, indelible, insolence) had the TL/SL transformation disrupting in a consonant cluster. These items were not included in the analysis.

Acknowledgements

Disclosure statement

Supplementary material

Supplementary material is available at journals.sagepub.com/doi/suppl/10.1080/17470218.2016.1271444.

Funding

References

Acha, J., & Perea, M. (2010). On the role of consonants and vowels in visual-word processing: Evidence with a letter search paradigm. *Language and Cognitive Processes*, *25*(3), 423–438.

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

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

Balota, D. A., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., & Treiman, R. (2002). *The English lexicon project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords*. Retrieved from http://elexicon.wustl.edu/

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D. M., Maechler, M., & Bolker, B. (2013). *Lme4: Linear mixed-effects models using S4 classes*. R package version 1.1-5.

Brand, M., Giroux, I., Puijalon, C., & Rey, A. (2007). Syllable onsets are perceptual reading units. *Memory & Cognition*, *35*, 966–973.

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

Buchwald, A., & Rapp, B. (2006). Consonants and vowels in orthographic representations. *Cognitive Neuropsychology*, *23*, 308–337.

Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition*, *37*, 243–297.

Carreiras, M., Duñabeitia, J. A., & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cerebral Cortex*, *19*, 2659–2670.

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

Chetail, F., & Content, A. (2013). Segmentation of written words in French. *Language and Speech*, *56*, 125–144.

Chetail, F., Drabs, V., & Content, A. (2014). The role of consonant/vowel organization in perceptual discrimination. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 938–961.

Chetail, F., Treiman, R., & Content, A. (2015). Effect of consonant/vowel letter organisation on the syllable counting task: Evidence from English. *Journal of Cognitive Psychology*, *5911*, 1–12.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance, VI* (pp. 535–555). Hillsdale: Erlbaum.

Coltheart, M., Rastle, K., Perry, C., Ziegler, J., & Langdon, R. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.

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

Duñabeitia, J. A., & Carreiras, M. (2011). The relative position priming effect depends on whether letters are vowels or consonants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1143–1163.

Fischer-Baum, S., Charny, J., & McCloskey, M. (2011). Both-edges representation of letter position in reading. *Psychonomic Bulletin & Review*, *18*(6), 1083–1089.

Forde, E. M. E., & Humphreys, G. W. (2005). Is oral spelling recognition dependent on reading or spelling systems? dissociative evidence from two single case studies. *Cognitive Neuropsychology*, *22*(2), 169–181. doi:10.1080/02643290442000040

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

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

Grainger, J., Granier, J. P., Farioli, F., van Assche, E., & van Heuven, W. J. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 865–884.

Grainger, J., & Hannagan, T. (2014). The architecture of writing systems *Written Language & Literacy*, *17*(2), 225–252. doi:10.1075/wll.17.2.03gra

Grainger, J., & van Heuven, W. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *The mental lexicon* (pp. 1–23). New York: Nova Science.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Oxford University Press, Clarendon Press.

Jones, M. N., & Mewhort, D. J. K. (2004). Case-sensitive letter and bigram frequency counts from large-scale English corpora. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 388–396. doi:10.3758/BF03195586

Kinoshita, S., & Norris, D. (2009). Transposed-letter priming of pre-lexical orthographic representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1–18.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package): R package version 2.0-11*. Retrieved from http://CRAN.R-project.org/package=lmerTest

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

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. *An account of basic findings. Psychological Review*, *88*, 375–407.

McCloskey, M., Badecker, W., Goodman-Schulman, R. A., & Aliminosa, D. (1994). The structure of graphemic representations in spelling: Evidence from a case of acquired dysgraphia. *Cognitive Neuropsychology*, *11*(3), 341–392. doi:10.1080/02643299408251979

McCloskey, M., Fischer-Baum, S., & Schubert, T. (2013). Representation of letter position in single-word reading: Evidence from acquired dyslexia. *Cognitive Neuropsychology*, *30*(6), 396–428. doi:10.1080/02643294.2014.880675

Miceli, G., Capasso, R., Benvegnù, B., & Caramazza, A. (2004). The categorical distinction of vowel and consonant representations: Evidence from dysgraphia. *Neurocase*, *10*(2), 109–121. doi:10.1080/13554790409609942

Morey, R. D., & Rouder, J. N. (2013). *BayesFactor: Computation of Bayes factors for common designs*. R package version 0.9.7. Retrieved from http://CRAN.R-project.org/package=BayesFactor

New, B., Araújo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science*, *19*, 1223–1227.

New, B., & Nazzi, T. (2014). The time course of consonant and vowel processing during word recognition. *Language, Cognition and Neuroscience*, *29*(2), 147–157.

Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General*, *137*, 434–455.

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.

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., & Acha, J. (2009). Does letter position coding depend on consonant/vowel status? Evidence with the masked priming technique. *Acta Psychologica*, *130*, 127–137.

Perea, M., & Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 97–120). New York: Psychology Press.

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.

R Core Development Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Version 3.0.3. Retrieved from http://www.R-project.org/

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.

Schubert, T. M., & McCloskey, M. (2015). Recognition of oral spelling is diagnostic of the central reading processes. *Cognitive Neuropsychology*, *32*(2), 80–88. doi:10.1080/02643294.2015.1031738

Tainturier, M.-J., & Rapp, B. C. (2003). Is a single graphemic buffer used in reading and spelling? *Aphasiology*, *17*(6-7), 537–562.

Vergara-Martínez, M., Perea, M., Marín, A., & Carreiras, M. (2011). The processing of consonants and vowels during letter identity and letter position assignment in visual-word recognition: An ERP study. *Brain and Language*, *118*, 105–117.

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., & Marton, Y. (2013). *The SERIOL2 model of orthographic processing* (pp. 1–76).

Yarkoni, T., Balota, D., & Yap, M. J. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979.