Chetail, F., Balota, D., Treiman, R., & Content, A. (accepted). What can megastudies tell us about the orthographic structure of English words? *Quarterly Journal of Experimental Psychology* (special issue on megastudies).

What can megastudies tell us about the orthographic structure of English words?

Fabienne Chetail[1], David Balota[2], Rebecca Treiman[2] & Alain Content[1]

[1]LCLD, CRCN, Université Libre de Bruxelles (ULB)

[2]Washington University in Saint Louis

Corresponding author:

Fabienne Chetail. Laboratoire Cognition Langage Développement (LCLD), Centre de Recherche Cognition et Neurosciences (CRCN), Université Libre de Bruxelles (ULB) – Av. F. Roosevelt, 50 / CP 191 - 1050 Brussels - Belgium.

E-mail address: fchetail@ulb.ac.be

Phone number: +32 2 650 26 07

Fax number: +32 2 650 22 09

Abstract

Although the majority of research in visual word recognition has targeted single syllable words, most words are polysyllabic. These words engender special challenges, one of which concerns their analysis into smaller units. According to a recent hypothesis, the organization of letters into groups of successive consonants (C) and vowels (V) constrains the orthographic structure of printed words. So far, evidence has been reported only in French with factorial studies of relatively small sets of items. In the present study, we performed regression analyses on corpora of megastudies (English and British Lexicon Project databases) to examine the influence of the CV pattern in English. We compared hiatus words, which present a mismatch between the number of syllables and the number of groups of adjacent vowel letters (e.g., *client*) to other words, controlling for standard lexical variables. In speeded pronunciation, hiatus words were processed more slowly than control words, and the effect was stronger in low-frequency words. In the lexical decision task, the interference effect of hiatus in low-frequency words was balanced by a facilitatory effect in high-frequency words. Taken together, the results support the hypothesis that the configuration of consonant and vowel letters influences the processing of polysyllabic words in English.

Keywords: megastudies, CV pattern, hiatus words, visual word recognition

The study of visual word recognition has been grounded on the study of short, monosyllabic words. Although this body of knowledge is a fundamental step to approach the complexity of reading processes, it may not fully generalize to polysyllabic words. A major challenge for polysyllabic words concerns how they are analysed into smaller units, explaining why modelling of polysyllabic word recognition lags behind monosyllabic word processing despite several decades of work (Perry, Ziegler, & Zorzi, 2010). Implementing a parsing process requires one to define the kind of units that determine the structure of words and the processes by which these units are perceived within a letter string.

Many different types of units have been proposed to influence word processing (see Patterson & Morton, 1985; Taft, 1991). Empirical evidence has been reported for graphemes (written representation of phonemes, e.g., Coltheart, 1978; Marinus & de Jong, 2011; Peereman, Brand, & Rey, 2006), onset/rime units (sub-constituents of syllables, e.g., Cutler, Butterfield, & Williams, 1987; Treiman, 1986; Treiman & Chafetz, 1987), graphosyllables (written representation of syllables, e.g., Spoehr & Smith, 1973; Carreiras, Alvarez & de Vega, 1993), morphemes (e.g., Longtin, Segui, & Hallé, 2003; Muncer, Knight, & Adams, in press), and BOSS units (basic orthographic syllabic structure, corresponding to a graphosyllable plus one or more consonants, e.g., Taft, 1979; Taft, Alvarez, & Carreiras, 2007).

The issue of reading units in polysyllabic letter strings has been approached from both phonological and orthographic perspectives. According to phonological views, the structure of letter strings is constrained by print-to-speech mapping, so that units within written words map onto linguistic units (e.g., Chetail & Mathey, 2009; Coltheart, 1978; Maïonchi-Pino, Magnan, & Ecalle, 2010; Spoehr & Smith, 1973). According to orthographic views, written word structure emerges from knowledge of letter co-occurrence regularities acquired through print exposure (e.g., Gibson, 1965; Prinzmetal, Treiman, & Rho, 1986; Seidenberg, 1987), and is therefore not necessarily isomorphic with spoken word structure. In line with the latter view, recent evidence in French supports the hypothesis that the organization of letters within words into consonants and vowels determines the perceived orthographic structure of letter strings (Chetail & Content, 2012, 2013, 2014). The way consonant and vowel letters are arranged is referred to as the "*CV pattern*", which can be viewed as the recoding of a letter string into a series of C and V category symbols (e.g., the CV pattern of *client* is CCVVCC). It is the status of the letters as consonants or vowels which is considered here, not the phonemes (see also Caramazza & Miceli, 1990). In the present study, we examine to what extent the CV pattern influences visual word recognition in English, using corpora of megastudies.

Role of the CV Pattern in Visual Word Processing

Chetail and Content (2012, 2013) demonstrated that the CV pattern of words constrains the orthographic structure of letter strings in French, with each vowel or series of adjacent vowel letters (henceforth, vowel cluster) determining one perceptual unit[1]. To ensure that the structure emerging from the organization of vowels and consonants is orthographic in nature, they used polysyllabic words for which the number of vowel clusters differed from the number of phonological syllables. In most words, groups of adjacent vowel letters map onto single phonemes (e.g., *people* - /piːpəl/, *evasion* - /ɪveɪʒən/) so that the number of syllables exactly matches the number of vowel clusters. However, this is not the case for words with a hiatus pattern, that is two (or more) adjacent vowel letters that map onto two phonemes (e.g., *oasis* - /əʊeɪsɪs/, *chaos* - /keɪɒs/, *reunion* - /rijunjən/). Hiatus words thus have one vowel cluster less than the number of syllables (e.g., *oasis* has a CV pattern with two vowel clusters, VVCVC, but it has three syllables /əʊ.eɪ.sɪs/). Although the proportion of hiatus words in languages is usually low (perhaps because articulation is optimal when there are consonants between vowels, e.g., Vallée, Rousset, & Boë, 2001), the mismatch between orthographic units (i.e., units based on vowel clusters) and phonological units (i.e., syllables) in hiatus words enables one to test whether the CV pattern determines the orthographic structure of words and influences visual word recognition.

Chetail and Content (2012) showed that French readers were slower and less accurate to count the number of syllables in written hiatus words such as *client* (/klijã/, two syllables, one vowel cluster) than in control words such as *flacon* (/flakõ/, two syllables, two vowel clusters). In hiatus words, erroneous responses most often corresponded to the number of vowel clusters. For example, participants were more likely to respond that *client* had one syllable than that it had three. If the structure of written words directly derived from their phonological form, no difference should have been found between control and hiatus words since both have the same number of spoken syllables. The finding that the number of units in hiatus words was underestimated suggests that letter strings are rapidly structured into units based on the CV pattern. The less efficient performance for hiatus words reflects the conflict between the perceptual orthographic structure derived from the distribution of vowel and consonant letters and the phonological syllabic structure.

Chetail and Content (2012) also examined the influence of the CV pattern in naming and lexical decision tasks to assess the extent to which letter organization affects word recognition. In the naming task, latencies were delayed for words exhibiting one vowel cluster less than the number of syllables, and this was especially true for four-syllable words. This delay was interpreted as reflecting the structural mismatch between the orthographic word form (e.g., *calendrier*, three units) and the phonological word form to be produced (e.g., /ka.lã.dʀi.je/, four syllables). In the lexical decision task, the direction of the effect varied as a function of word length. Trisyllabic hiatus words (e.g., *sanglier*, /sã.gli.je/) tended to be recognized more rapidly than control words (e.g., *saladier*, /sa.la.dje/), while the effect was reversed for words with four syllables. The facilitatory effect of hiatus for the shorter words was explained in terms of sequential processing, words with fewer orthographic units being more quickly processed because they would need fewer steps (see Ans, Carbonnel, & Valdois, 1998; Carreiras, Ferrand, Grainger, & Perea, 2005). For longer words, the lexical identification process takes more time, increasing the likelihood that phonological assembly processes noticeably influence performance as in the naming task and yielding a net inhibitory effect. Based on these results, Chetail and Content concluded that the processing of polysyllabic words may engage a level of orthographic representations based on the CV pattern. The orthographic structure computed at this level can cause interference at later levels of processing when it does not match the phonological structure.

Cross-Linguistic Investigation

To date, the influence of CV letter organization on word processing has been examined only in French, and in factorial studies with limited sets of words. To ensure that CV pattern effects are not language-specific, they need to be examined in other languages. Orthographies vary in the number and complexity of graphemes, their mapping onto phonemes, and the syllabic structure of the corresponding spoken language (see Seymour, Aro, & Erskine, 2003; van den Bosch, Content, Daelemans, & de Gelder, 1994), and one might wonder whether the influence of the CV structure is present despite these differences. A body of evidence indicates that language-specific characteristics constrain the nature of reading units. For example, the importance of rimes in English may relate to the fact that the consistency of the grapho-phonological mapping is stronger for these units than for phonemes (e.g., Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995; De Cara & Goswami, 2002). The salient role of graphosyllables in Spanish could be due to the relatively simple syllabic structure of this language (Alvarez , Carreiras, & de Vega, 2000). Finally, the fact that evidence for the BOSS unit was reported mainly in English (e.g., Taft, 1979, 1987, 1992, but see Rouibah & Taft, 2001) could reflect the relatively high proportion of closed syllables in English. Although the nature of reading units may vary across languages, a common process could underlie their extraction. Units like the rime, the graphosyllable, and the BOSS share the characteristic of being centred on one or several vowel letters, preceded or followed by consonant clusters. Therefore, an early parsing based on the CV pattern may lead to coarse orthographic units that are refined according to the specific characteristics of an orthography. Supporting this hypothesis requires one to test the influence of the CV pattern in languages with different orthographic characteristics than French. The present study is the first to examine CV structure effects in English.

A further reason to examine the effect of the CV pattern in English is that it permits to investigate the relation between CV parsing and graphemic parsing. Graphemic units have sometimes been considered as the basis of perceptual representations (e.g., Rastle & Coltheart, 1998; Rey, Ziegler, & Jacobs, 2000). Indeed, a graphemic parsing stage has been implemented in the sublexical route of influential recent word recognition models (e.g., Perry, Ziegler, & Zorzi, 2007, 2010). Letter strings handled by the non-lexical route are parsed into graphemes, which are then assigned to onset, nucleus and coda constituent slots, so that graphosyllables are aligned with phonology. One way to investigate the relation between CV parsing and graphemic parsing is to examine whether processing differences between hiatus and control words vary according to the probability that the bigram marking the hiatus maps onto one or two phonemes. Some English vowel clusters (e.g., *eo*) map onto two phonemes in certain words (e.g., *video*) and one phoneme in other words (e.g., *people*). In the former case, the word entails a hiatus, and the mismatch between the number of orthographic units as defined by the CV pattern and the number of syllables should impair naming (Chetail & Content, 2012). In the latter case, with words like *people*, no mismatch is present. This mapping ambiguity does not exist in French (e.g., *ao* as in *chaos* systematically maps onto two phonemes), but it is a frequent feature in English. Furthermore, some vowel groups map onto a single phoneme more often whereas other most frequently correspond to two phonemes (see table 7). We will refer to this as the degree of graphemic cohesion of the vowel cluster. In naming, if CV parsing occurs before graphemic parsing, the hiatus effect should be stronger when the bigram frequently maps onto a single grapheme (e.g., *ai*, strong graphemic cohesion) than when the bigram frequently maps onto two graphemes (e.g., *iu*, weak graphemic cohesion), because more time may be necessary to break the bigram and to assign each of its component to a different grapheme slot. On the contrary, when the phonological structure of words is not predominantly activated, as in the lexical decision task (see e.g., Balota et al., 2004; Grainger & Ferrand, 1996), the hiatus effect should not be modified as a function of graphemic cohesion.

The Present Study

The aim of the present study was to examine the effect of the CV pattern on visual word recognition in English, operationalized by the presence or not of a hiatus pattern within words. We first conducted item-based regression analyses on naming and lexical decision latencies from the English Lexicon Project (ELP, Balota et al., 2007) contrasting hiatus and non-hiatus words. The ELP is the first behavioral database for a very large number of words (~40,000), and collected on native speakers of American English. This has been followed by the creation of similar large-scale databases in other languages (e.g., Ferrand et al., 2010, in French; Keuleers, Diependaele, & Brysbaert, 2010, in Dutch; Keuleers, Lacey, Rastle, & Brysbaert, 2012, in British English). These corpora are usually analysed with regression methods, which offer a number of advantages compared to standard factorial studies (e.g., no need to orthogonally manipulate factors of interest, better control of other variables), although most specialists also argue that the two approaches should be used together.

To assess the impact of CV letter organization on the processing of English words, we conducted multiple regression analyses on naming and lexical decision latencies including the word type factor (hiatus vs. non-hiatus), together with standard variables known to affect latencies, namely word frequency, number of letters, bigram frequency, orthographic neighborhood, number of syllables, and consistency (see, e.g., Balota et al., 2004; New, Ferrand, Pallier, & Brysbaert, 2006; Treiman et al., 1995; Yap & Balota, 2009). Word frequency explains the largest part of variance in visual word processing, with words of high frequency being processed more rapidly (e.g., Balota et al., 2004; New et al., 2006; Yap & Balota, 2009). Over and above word frequency, other factors influence processing latencies. Orthographic typicality, as measured by neighborhood density (i.e., number of words with an orthographic form similar to the target word) or bigram frequency (e.g., summed frequency of bigrams within words), usually facilitates word processing (e.g., Andrews, 1997; Massaro, Venezky, Taylor, 1979). Additionally, words with regular letter-to-sound correspondences (i.e., feedforward consistency) or sound-to-letter correspondences (i.e., feedback consistency) are processed more rapidly than inconsistent words (e.g., Stone & Van Orden, 1994; Yap & Balota, 2009; but see Kessler, Treiman, & Mullennix, 2008 for questions about the role of feedback inconsistency). Finally, length of a word in number of letters or syllables has been reported to affect word processing (e.g., Balota et al., 2004; Ferrand & New, 2003; Muncer & Knight, 2012).

Our prediction was that, if the perceptual structure of written words is determined by the CV pattern in English, we should find an effect of the hiatus when the influence of the other lexical variables is partialled out. In naming, we expected hiatus words to be processed more slowly than control words. We also expected the effect to be modulated by the frequency with which the vowel cluster marking the hiatus is associated with one phoneme. We expected the hiatus effect to be weaker or even non-existent in the lexical decision task, due to the balance between a facilitatory effect for short or high-frequency words and an inhibitory effect for long or low-frequency words, replicating the pattern obtained by Chetail and Content (2012). To test these hypotheses, we relied on the two corpora available from megastudies in English (the English Lexicon Project, ELP, Balota et al., 2007, and the British Lexicon Project, BLP, Keuleers, Lacey, Rastle, & Brysbaert, 2012.

Method

The main analysis aimed at testing the hiatus effect in English with the ELP database (Balota et al., 2007). The ELP contains 40,481 words with behavioral data collected in lexical decision and naming tasks. For each entry, lexical characteristics of words are also included (e.g., word frequency, number of letters, morphemic parsing). The phonological transcription of word pronunciation is based on General American standard, and largely comes from the Unisyn Lexicon developed by the Centre for Speech Technology Research (Center for Speech Technology Research, University of Edinburgh, n.d.).

A preliminary step was to identify orthographic hiatus words, that is, words entailing two adjacent vowel letters mapping onto two distinct phonemes (e.g., *chaos*, *oasis*, *lion*). We used an algorithm that detected entries including both a cluster of two adjacent vowel letters in the orthographic word form (i.e., A, E, I, O, U, Y) and a cluster of two adjacent vowel phonemes (or diphthongs) in the phonological form (i.e., /ɑ/, /æ/, /ɜ/, /ə/, /e/, /i/, /ɪ/, /ɔ/, /ʌ/, /ʊ/, /u/, /ɑɪ/, /ɑʊ/, /eɪ/, /oɪ/, /əʊ/). A manual check enabled us to exclude from the set of hiatus words entries for which the two clusters were not aligned. In addition, words in which the phonological hiatus pattern mapped onto two vowel clusters (*dewy* - /dui/) were treated as non-hiatus words (see Chetail & Content, 2012). Finally, words with the letter Y deserved special attention because Y can act as a consonant (e.g., *rayon*) or a vowel (e.g., *cycle*). Words with a Y surrounded by two vowels (e.g., *rayon*) were considered non-hiatus words whereas words with a Y only preceded or followed by a vowel cluster (e.g., *dryad*) were considered hiatus words. This procedure led us to identify 2,469 hiatus words out of 40,481, 6.01% of the set.

To perform the regression analyses, we restricted the set to words that were polysyllabic (N = 34,186), monomorphemic (N = 6,025) and which had no missing data (N = 5,678). In the final set of 5,678 words, 400 had a hiatus pattern (7.04%)[2]. A dummy variable was used to code for word type (hiatus vs. control words). In addition to this predictor, the following variables were used as predictors:

Word frequency: Measure of lexical frequency (log transformed) provided in the ELP and based on film subtitles (*LgSUBTLWF*, see Brysbaert & New, 2009; Keuleers, Brysbaert & New, 2010).

Number of letters: Letter count in orthographic word forms, from the ELP.

Orthographic neighborhood density: Orthographic Levenshtein distance (*OLD*), provided in the ELP. Levenshtein distance represents the number of letter insertions, deletions, and substitutions needed to convert one string to another one. For example, the distance between *pain* and *train* is 2: substitution of P by R, and insertion of T. The specific measure that we used was the mean OLD computed on the 20 closest neighbors of each entry (see Yarkoni, Balota, & Yap, 2008).

Bigram frequency: Average bigram count for each word (token count), provided in the ELP.

Number of syllables: Number of syllables in phonological word forms, provided in the ELP. For some hiatus words, we adjusted the count because words were coded with one syllable less than in their standard American pronunciation (e.g., *lion* was coded as a monosyllabic word but we counted it as two syllables, see Merriam-Webster online).

Consistency: Four measures of consistency from Yap (2007): FFO consistency, FFR consistency, FBO consistency, and FBR consistency. These measures reflect either feedforward consistency (FF, spelling-to-sound) or feedback consistency (FB, sound-to-spelling), computed either on rimes (R) or on onsets (O), across syllabic positions.

First phoneme identity: When relevant for the analyses, we used 13 dummy variables to code for first phoneme properties (see Treiman et al., 1995; Yap & Balota, 2009).

The four dependent variables were the standardized mean reaction time and mean accuracy in the two tasks. As pointed out by Balota et al. (2007), a *z*-score transformation on the raw reaction times minimizes the influence of a given participant's processing speed and variability, making it possible to directly compare performance on different words (see Faust, Balota, Spieler, & Ferraro, 1999). The descriptive statistics for the predictors and dependent variables are provided in Table 1 and the inter-correlations in Table 2. In the regression analyses, all continuous predictors were centred around their mean (see Cohen, Cohen, West, & Aiken, 2003).

Results

First, we ran global regression analyses on the whole word set (N = 5,678) to test the effect of word type. Then, we tested the reliability of this effect in two additional analyses. The first used the Monte Carlo method to run multiple tests with different random samples of words, and the second used another English database, the British Lexicon Project (BLP, Keuleers et al., 2012). Finally, we examined whether the effect of word type varied according to graphemic cohesion.

Global Regression Analyses on ELP

Three-step hierarchical regression analyses were carried out on both lexical decision and naming performance. In the first step, we entered all the first phoneme predictors and the lexical predictors (word frequency, number of letters, number of syllables, mean bigram frequency, OLD, and the four measures of consistency). These variables were entered in a single step because we were not interested in their distinct roles but we wanted to ensure that they did not influence the results. In the second step, we entered word type (hiatus vs. control). In the last step, the interaction between word type and word frequency was added. This procedure enabled us to assess the effect related to the presence of hiatus when the effect of standard lexical variables has been partialled out (step 2) and to examine to what extent the word type effect varied with word frequency (step 3). The results are presented in Table 3.

The first phoneme and lexical variables accounted for a large portion of the variance in the first step (between 55% and 58% for reaction times, and between 30% and 44% for accuracy). When all these variables were controlled, adding the word type predictor led to a model that significantly accounted for more variance than the step 1 model, but only for naming (increase of 0.1% in reaction times). The positive correlation between word type and RTs showed that words with an orthographic hiatus pattern were processed more slowly than control words. Accuracy tended to be lower for hiatus words. Introducing the interaction between word type and word frequency at step 3 led to a significant R^2 improvement in both tasks. To interpret the interaction in naming, we plotted the estimated performance in Step 2 as a function of word frequency for each type of word. As shown in Figure 1, hiatus words were processed more slowly than control words in naming when they were of low frequency (left panel), and the effect disappeared and even reversed for high-frequency words. In lexical decision (right panel), the presence of an interaction without a main effect of word type suggests that, for the whole word set, the inhibitory effect of word type for the least frequent words and the facilitatory effect for the most frequent words balance each other. On the contrary, the fact that the main effect of word type remained significant in Step 3 in the naming RTs shows that there is a genuine overall trend towards interference in this task when words include a hiatus.

Additionally, step 3 was run with the interaction between word type and number of letters instead of the interaction between word type and word frequency. Including this interaction did not lead to an additional significant R^2 improvement nor to a significant contribution of the interaction (LDT RT: β = .005, *p* = .51; LDT Accuracy: β = -.007, *p* = .15; Naming RT: β = -.007, *p* = .35; Naming Accuracy: β = -.006, *p* = .07). The interaction was still not significant when the number of syllables was considered instead of the number of letters.

Monte Carlo method for multiple tests on the ELP

There was a large difference in the number of observations for the two levels of the critical predictor (hiatus vs. control), and one could argue that using such unbalanced sets artificially increases the statistical power of the comparison. We therefore used the Monte Carlo method to run multiple tests using different random samples of control words. Hence, we conducted regression analyses on the ELP with 800 observations (rather than 5,678). Half corresponded to the 400 hiatus words and the other half corresponded to 400 control words randomly selected among the full set of 5,278 words. We ran 100 regressions for each of the four dependent variables, each regression being therefore performed with a different random selection of control words (see also Keuleers et al., 2010; Keuleers et al., 2012, for statistical tests on multiple random samples). The analyses were identical to those described in the previous section, except that the first phoneme variables were not included because the complete crossing of levels of the 13 first phoneme variables was not represented in each random draw. Continuous variables were centred on the 800 items at each draw. Table 4 presents the mean RT and accuracy coefficients for the effects of word type (main effect, interaction with word frequency), and the mean variance explained by the model at each step across the multiple draws. Figure 2 shows the distribution of the regression draws as a function of level of significance of word type effects on reaction time and accuracy in naming and lexical decision.

The interaction between word type and word frequency was significant (alpha = .05) in more than 95% of the regressions in both tasks (Table 4) and the level of significance was higher in lexical decision than in naming (Figure 2). In addition, the main effect of word type was significant in 100% versus 1% of the regressions on reaction times in naming and lexical decision, respectively (Figure 2). This confirms the results obtained in the general regression analysis. The presence of a hiatus pattern within letter strings genuinely influences written word processing, depending on word frequency. At a methodological level, the results show that the effects previously found were not due to an unbalanced number of observations for the two types of words. This further exemplifies the importance of megastudies in affording a large number of repeated tests on different word samples, which provides a powerful and flexible method to test hypotheses.

British Lexicon Project (BLP)

A second way to test the reliability of the hiatus effect was to replicate the results on the BLP (Keuleers et al., 2012). The BLP is a database of lexical decision times for English mono- and bisyllabic words and nonwords collected with British participants, containing 28,594 items. As in our analyses with the ELP, we selected only monomorphemic words without missing data. This led to a set of 4,398 bisyllabic words, which included 4,340 control items and 58 hiatus words[2]. The smaller number of items compared to ELP (5,278 control and 400 hiatus words) can be explained by the fact that first, the BLP tested bisyllabic words but no words with more than two syllables, and second the British pronunciation of bisyllabic words leads to fewer hiatus than in American English. We conducted the same regression analysis as previously, except that neither the number of syllables (only bisyllabic words) nor consistency measures (not available for the BLP) were included in the model. The frequency used was the Zipf measure of the frequency counts in British subtitles (van Heuven, Mandera, Keuleers, & Brysbaert, 2014). As shown in Table 5, the general regression analysis replicated the effects on the ELP. No significant main effect of word type was observed, but, as previously, the interaction between word type and word frequency was significant.

As with the ELP data, we conducted multiple tests using different balanced sets of hiatus and control words selected among the 4,398 items of the BLP. Table 6 presents the mean RT and accuracy coefficients for the effects involving word type, and the mean variance explained by the model at each step, across the multiple draws. Figure 3 shows the distribution of the 100 regressions as a function of level of significance of word type effects on regressed reaction times. The results corroborate those of the previous analyses with the ELP dataset. The word type effect was not significant in the LDT reaction times or on accuracy estimates, while the interaction between word type and word frequency was significant (alpha = .05) in 24% of the draws in the reaction times and 64% in accuracy. The relative weakness of the effects with respect to the ELP may be attributed to the fact that 58 pairs were contrasted in the analysis based on the ELP versus 400 in the ELP.

Hiatus Effect and Graphemic Cohesion

In the last analysis, we tested the extent to which the hiatus effect varied according to the frequency that a critical letter bigram (e.g., *ea*, *eo*) corresponds to two adjacent graphemes and maps onto two phonemes (hiatus words, e.g., *create*, *video*), or corresponds to a single grapheme and maps onto one phoneme or a diphthong (e.g., *season, people*). We selected 21 bigrams which can code for hiatus patterns in English words[3]. For each bigram, we computed three measures based on the full ELP word set (Table 7): the token frequency of the mapping with a single grapheme (i.e., summed frequency of all the words that contain the grapheme), the token frequency of the mapping with two graphemes (hiatus words), and graphemic cohesion, corresponding to the ratio of the former to the sum of the former and the latter (i.e., frequency of the bigram on all words including the given bigram as a grapheme or as a hiatus). To isolate the relevant graphemes, we relied on the English grapheme-to-phoneme correspondences provided by Lange (2000), the only exception being for the bigram UO. According to Lange (2000), UO is not a grapheme in English, but we found that it was coded as such in the ELP (e.g., *fluorescent* -> /flʊresənt/). Graphemic cohesion ranges from 0 to 1, with values close to one reflecting a bigram that most often maps onto a single phoneme and that rarely occurs in hiatus words. As Table 7 shows, the distribution of graphemic cohesion over the 21 bigrams is bimodal. The bigrams either mostly map onto one grapheme (nine bigrams with a graphemic cohesion superior to 0.80) or onto two graphemes (eight bigrams with a graphemic cohesion inferior to 0.20).

To analyse the effect of graphemic cohesion, we used the ELP and we first carried out a regression analysis on the set of hiatus words only, including graphemic cohesion as an additional continuous predictor at the second step. Out of the 400 initial hiatus words, only the 377 words that included one and only one critical bigram were included in the analysis. In the naming task, there was a marginal effect of graphemic cohesion on RTs (β = .073, *p* = .07) in the predicted direction, such that hiatus words containing vowel bigrams of high graphemic cohesion gave rise to longer naming RTs. The effect failed to reach significance for accuracy (β = -.020, *p* = .17) and was not present at all in the lexical decision task (RTs: β = .002, *p* = .95, Accuracy: β = -.026, *p* = .26).

To further examine the potential role of graphemic cohesion, we conducted a second analysis that included a comparison with control words. Hiatus words were separated into two groups according to graphemic cohesion, weak (graphemic cohesion = 0, N = 234) or strong (graphemic cohesion > 0.50, N = 133). We conducted 100 regressions for both groups, each being performed with a set of 234 and 133 control words, respectively, and control words being randomly selected among the 5,278 control words. We performed the same regression analyses as in the general analysis (step 1: lexical variables, step 2: lexical variables, word type, step 3: lexical variables, word type, word type x frequency). If graphemic cohesion influenced the hiatus effect, the hiatus effect should be stronger in the strong cohesion set than in the weak cohesion set. As can be seen in Figure 4, this pattern was found in naming but not in lexical decision. At step 3, 91% of the regressions led to a significant hiatus effect for strong graphemic cohesion versus 64% for weak graphemic cohesion in the RTs analyses (18% vs. 1% in the accuracy analyses), whereas there was no difference in the lexical decision (4% vs. 1% in reaction times, and 2% vs. 0% in accuracy for strong and weak graphemic cohesion respectively).

Discussion

The aim of the present study was to examine the effect of hiatus –a marker of the role of the CV pattern– on visual word recognition in English. To do so, we used a regression method to contrast naming and lexical decision latency and accuracy from English corpora of megastudies for hiatus words and non-hiatus words. After removing the effect of variables known to influence visual word recognition, we found a reliable effect of word type in the naming task, with words entailing a hiatus pattern being processed more slowly than control words. The effect was stronger for low-frequency words than high-frequency ones. In the lexical decision task, although there was no significant main effect for hiatus, the interaction between word type and word frequency showed a similar interference effect of hiatus in low-frequency words as in naming, counterbalanced by a facilitatory effect of hiatus in high-frequency words. This pattern of results was reproduced across hundreds of regressions on random samples of words and was very similar with two different English databases (in the lexical decision task). A final analysis showed that the word type effect in naming was stronger when the bigram coding for the hiatus corresponds to a single grapheme in most words, although it was significant even when the bigram never corresponds to a single grapheme.

CV Pattern and Orthographies

The respective role of consonants and vowels in visual word recognition has been an issue of major interest over the last decades, and it has been approached from different perspectives. First, Berent and Perfetti (1995) proposed that the phonological conversion of consonants occurs faster than that of vowels (see also Marom & Berent, 2010). Although the hypothesis was supported by evidence from English, the two-cycles hypothesis has not been confirmed in more transparent orthographies (e.g., Colombo, Zorzi, Cubelli, & Brivio, 2003), suggesting that it may be dependent on the differential consistency of vowels and consonants in a given language. Second, studies disturbing consonant or vowel information by selective transposition or deletion suggest that consonants provide stronger constraints on lexical selection than vowels (e.g., Duñabeitia & Carreiras, 2011; Lupker, Perea, & Davis, 2008; New, Araújo, & Nazzi, 2008; Perea & Acha, 2009). Third, the present findings, together with other recent studies (e.g., Chetail & Content, 2012, 2013, 2014), support the psychological reality of orthographic units mediating visual word recognition which are determined by the arrangement of consonant and vowel letters.

In the naming task, the interference associated with the presence of a hiatus, previously found in French (Chetail & Content, 2012), is extended here to English. This effect can be accounted for by a conflict between two levels of representation, as suggested by the findings in the syllable counting task in French. Orthographically, a hiatus word like *video* is structured into two units because it contains two vowel clusters, *i* and *eo.* At a phonological level, however, this word is structured into three syllables. The orthographic structure is salient at first. However, in the word naming task, the need to produce an oral response involves the activation of the syllabic structure during phonetic encoding and articulatory preparation (e.g., Levelt & Wheeldon, 1994). The delay in naming, we propose, ensues from the mismatch between the two representations. Given that low-frequency words are processed more slowly than high-frequency words, the conflict between the two representations would have more time to influence processing, causing a stronger word type effect in low-frequency words.

In the LDT, the reliable interaction between word type and word frequency shows that hiatus words are processed more efficiently than control words when they are of high frequency, whereas the opposite is found when they are of low frequency. Assuming that phonology plays a more important role with low frequency words, the interference effect is consistent with the results obtained in the naming task and may reflect a conflict between the orthographic structure and the phonological structure. When words are of high frequency, processing is more orthographically oriented (i.e., phonological recoding may be less influential). In that case, according to the hypothesis of sequential processing of orthographic information (see Ans et al., 1998; Carreiras et al., 2005), hiatus words are identified more rapidly than control words because they have fewer orthographic units. This explanation was devised to account for the length by word type interaction in French (Chetail & Content, 2012). Here, however, we found no hint of such an interaction in English. Although this issue deserves further attention, a potential explanation is that English words are on average shorter than French words. This reduces the probability of detecting a crossover interaction between number of letters and word type using the English megastudies. Actually, the three- and four-syllable words in Chetail and Content (2012) were 8.24 and 10.13 letters long respectively, whereas the length was 7.34 and 8.95 letters in the corresponding set from ELP.

One could wonder whether the effects found with hiatus words can be explained by phonological variables, especially in naming, since articulation may be less optimal when there is no consonant between vowels and hiatus structures are relatively infrequent. However, there is direct evidence in French that the hiatus effect is not driven by phonological or production characteristics. Chetail and Content (2012) compared two kinds of hiatus words, both exhibiting two contiguous phonological full vowels. In one case, the phonological hiatus corresponded to adjacent vowel letters (e.g., *chaos*, /ka.o/) thus entailing an orthographic/phonological mismatch as in the present study, whereas in the other case, the phonological hiatus corresponded to two vowel letters separated by a silent consonant (e.g., *bahut*, /ba.y/), thus leading to two disjoint orthographic vowel clusters. In the latter words, although the phonological form contains two contiguous vowels, the alternation of orthographic consonants and vowels determines a segmentation that is consistent with the syllabification (i.e., two vowel clusters in bisyllabic words). Accordingly, orthographic hiatus words like *chaos* but not phonological hiatus words like *bahut*, were processed less efficiently than the control words.

In the present study, we provide the first evidence that CV pattern effects are not restricted to French, and generalize to English. The presence of a hiatus effect in two orthographies varying in the complexity of letter-to-sound mapping and syllabic complexity (e.g., Seymour et al., 2003; Ziegler et al., 1996; van den Bosch et al., 1994) could suggest that the basic structure of letter strings is determined by the CV pattern of words in any orthography with consonant and vowel letters. This parsing process leads to coarse orthographic units, centred on vowel clusters to which adjacent consonants aggregate. These units would be refined according to the specificities of the given orthography, such as the size of prototypical graphosyllables or the regularity of letter-to-sound mapping. In that sense, the CV pattern would not only provide a source of information to access reading units but also afford robust invariant cues that guide initial parsing. A parsing based on the CV pattern does not reflect a universal procedure, implicated in any writing system (see Frost, 2012), since it is relevant only in writing systems with vowel and consonant letters (not in Chinese for example), but it may provide a useful framework for understanding polysyllabic word parsing in any alphabetic orthography. Further examination of the CV pattern hypothesis will therefore be necessary in other languages.

CV Parsing and Graphemic Parsing

Another key result of the present study is the variation of the hiatus effect as a function of graphemic cohesion (see also Spinelli, Kandel, Guerassimovitch, & Ferrand, 2012, for effects of graphemic cohesion). As explained in the introduction, testing this interaction is not possible in French because bigrams marking hiatus only map onto two graphemes (e.g., *ao* always maps onto /aɔ/). In English, on the contrary, a bigram (e.g., *eo*) can map onto two phonemes (e.g., *vid<u>eo</u>*) or onto one phoneme (e.g., *p<u>eo</u>ple*), giving us the opportunity to examine this new issue. Importantly, megastudy databases are ideal for addressing this.

The results showed that, in the naming task, the hiatus effect tended to be stronger when the bigram coding the hiatus pattern frequently maps onto a single complex grapheme otherwise (i.e., strong graphemic cohesion). This outcome supports the hypothesis that parsing based on grapheme units might occur, but only after CV parsing. For hiatus words, the CV parsing is not compatible with the graphemic parsing, because the vowel cluster (*eo* in *video* for example) is part of one unit whereas it corresponds to two graphemes. Therefore, the graphemic segmentation process has to break the vowel cluster of hiatus words. In the CDP++ model for example, (Perry et al., 2010), “complex graphemes [are] preferred over simple ones whenever there is potential ambiguity” (p. 114). Vowel clusters that frequently map onto a single grapheme (i.e., strong graphemic cohesion) may therefore preferentially be assigned to a single grapheme slot, leading to erroneous combinations of graphosyllable constituents when the system attempts to generate the pronunciation. The failure to access the correct phonology would require at least another graphemic parsing attempt, delaying word pronunciation and resulting in a strong hiatus effect. On the contrary, when the vowel cluster is less strongly associated to a single grapheme (i.e., weak graphemic cohesion), the two vowel letters would likely be correctly assigned to two different graphemes if one assumes that graphemic parsing is sensitive to grapheme frequency. Thus, with this additional hypothesis, the modulation of the hiatus effect by graphemic cohesion could be explained. Importantly, the hiatus effect was present even when the vowel bigram never maps onto one grapheme (i.e., weak graphemic cohesion). This is consistent with findings in French for which hiatus bigrams never map onto a grapheme and unambiguously correspond to two graphemes (e.g., *éa* in *océan* systematically maps onto /eã/). The convergent observations in French and English demonstrate that the hiatus effect cannot be reduced to grapho-phonemic consistency effects, since it is present even when there is no print-to-sound mapping ambiguity.

The presence of an interaction between word type and graphemic cohesion only in the naming task suggests that graphemic cohesion influences processing only during print-to-sound mapping and challenges the idea that graphemes constitute perceptual units at an early level of visual word recognition. This conclusion is supported by a recent study by Lupker, Acha, Davis, and Perea (2012), who assessed the role of graphemes in visual word processing. They reasoned that, if graphemes are perceptual units, disturbing letters in a complex grapheme (e.g., *TH*) should produce a larger effect on word processing than when letters that constitute two graphemes are disturbed (e.g., *BL*). Using transposed-letter priming, they found no difference between the two conditions in a masked lexical decision study in either English or Spanish. Both *anhtem* and *emlbem* facilitated lexical decisions for the target words *ANTHEM* and *EMBLEM* respectively, compared to a control condition. This led the authors to conclude that multiletter graphemes are not perceptual units involved in early stages of visual word identification.

Implications for Models of Orthographic Encoding

The current results have direct implications for current models of orthographic encoding. In the last decade, much effort in the field of visual word recognition has been dedicated to producing models accounting for the early stages of visual analysis, letter identification, and letter position and sequence coding in multi-letter strings (see Frost, 2012, for a review). This research has been especially based on transposed-letter priming effects (e.g., *anwser* facilitates the processing of *ANSWER* as much as an identity prime, Forster, Davis, Schoknecht, & Carter, 1987) and superset/subset priming (e.g., *blck* facilitates the processing of *BLACK*, Peressotti & Grainger, 1999). Importantly, this work has led to abandon the hypothesis of strict positional coding which assumed separate slots for specific letter positions, and instead, different schemes offering positional flexibility have been proposed.

As suggested by Taft and colleagues (e.g., Lee & Taft, 2009, 2011; Taft & Krebs-Lazendic, 2013), one limitation of these models is that they assume that the only information that plays a role in early orthographic processing is the identity and absolute or relative position of the letters and bigrams. For example, in open-bigram models (e.g., Grainger & Van Heuven, 2003; Whitney, 2001), stimuli activate bigrams corresponding to adjacent and non-adjacent letters (e.g., FO, FR, FM, OR, OM, and RM for *FORM*, and FR, FO, FM, RO, RM, and OM for *FROM*). Due to the high overlap of activated bigrams (5/6 in the *FORM/FROM* example), a prime created by the transposition of two letters is as good as the base word itself. According to the spatial gradient hypothesis (Davis, 2010; Davis & Bowers, 2006), the orthographic representation depends on a specific pattern of activation of its component letters, with activation decreasing from left to right as a function of letter position within the string. Hence, in both *FORM* and *FROM*, the letters F and M are the most and the least activated, respectively, and O is more activated than R in *FORM* whereas R is more activated than O in *FROM*. Again, both letter strings are therefore coded by relatively similar patterns of letter activation. Finally, according to the noisy positional coding scheme (e.g., Gomez, Ratcliff, & Perea, 2008; Norris, Kinoshita, & van Casteren, 2010), the activation of each letter extends to adjacent positions, so that the representation of *FORM* is strongly activated by R in the third position but also by R in the second position. All of these models assume that the underlying structure of words is a plain string of letters or bigrams. However, the fact that transposed-letter effects are influenced by onset-coda structure (e.g., Taft & Krebs-Lazendic, 2013), morphological structure (e.g., Perea, abu Mallouh, & Carreiras, 2010; Velan & Frost, 2009), and the CV pattern (Chetail, Drabs, & Content, in press) argues for richer and more complex orthographic representations, beyond linear letter strings. The present study provides further support for this view.

One possibility would be to incorporate an intermediate level of orthographic representations based on vowel clusters (Chetail, Drabs, & Content, in press). Although the idea of an intermediate level of representations between letters and word form is far from new (e.g., Patterson & Morton, 1985; Shallice & McCarthy, 1985; Taft, 1991; Conrad, Tamm, Carreiras, & Jacobs, 2010), the specificity of the current proposal is that the grouping strictly ensues from orthographic characteristics, namely, the arrangement of consonant and vowel letters. In this view, a minimal perceptual hierarchy might include four levels of representation: features, letters, vowel-centered units (i.e., orthographic units based on the CV pattern of words), and orthographic word forms. Vowel-centered units would thus both serve to contact lexical representations and to encode the identity and spatial position of substrings from the sensory stimulation. Furthermore, the number of active vowel-centered nodes or the summed activity in that layer might provide a useful cue to string length and structure. This proposition is consistent with recent evidence showing that the number of vowel-centered units influences the perceived length of words (Chetail & Content, 2014), even with presentation durations so short that stimuli were not consistently identified. In this architecture, it can still be assumed that graphemes are extracted and serve as the basis for a separate phonological conversion procedure where graphemic units are inserted into a grapho-syllabic structure with onset, nucleus and coda slots (as in the CDP++ model, Perry et al., 2007, 2010). In this context, vowel-centered units might provide a clue to extract the grapho-syllabic and phonological structure, since vowel-centered units most of the time correspond to graphosyllables. One advantage of vowel-centered units would be to code the orthographic structure of letter strings according to a definite and fixed scheme, independent of ortho-phonological mapping inconsistencies.

Conclusion

Relying on the recent development of megastudies, we provided evidence in English for a new hypothesis according to which the configuration of consonant and vowel letters (i.e., the CV pattern) influences visual word processing. In line with previous studies with French, the results suggest that the CV pattern of words shapes perceptual representations at an early stage, before a potential graphemic parsing stage. Importantly, the information available from megastudies extends the hiatus effect across a large set of stimuli, affords control over potentially correlated variables via regression techniques, and permits extension to the novel variable of graphemic cohesion. The similarity of results in the previous experiments in French and in the present study in English provides convergent evidence for the importance of the CV pattern in languages with different orthographic characteristics. Clearly, this works needs to be extended to additional languages. Fortunately, the need for cross-linguistic comparisons is reflected in the ongoing development of megastudies in both alphabetic (e.g., Keuleers et al., 2010, in Dutch; Yap, Liow, Jalil, & Faizal, 2010, in Malay) and non-alphabetic (e.g., Sze, Liow, & Yap, 2014, in Chinese) languages.

Footnote

[1] Strictly speaking, a vowel cluster refers to a group of two vowel letters or more, but for the sake of simplicity the term will be used hereafter to refer to both single vowels and groups of vowels preceded and/or followed by consonants.

[2] The list of the hiatus words is available from the first author's home page (http://lcld.ulb.ac.be/lequipe/fabienne-chetail).

[3] Six bigrams were discarded from this analysis (e.g., AE, II, UU, YA, YE, YU) because they were present in very few words, of very low frequency.

References

Alvarez, C. J., Carreiras, M., & De Vega, M. (2000). Syllable-frequency effect in visual word recognition: Evidence of sequential-type processing. *Psicologica*, *21*, 341–374.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4*, 439–461. doi:10.3758/BF03214334

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, *105*, 678-723. doi:10.1037/0033-295X.105.4.678-723

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*, 283–316.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445-459. doi:10.3758/BF03193014

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*, 977-990. doi:10.3758/BRM.41.4.977

Carreiras, M., Alvarez, C. J., & De Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*, *32*, 766-780. doi:10.1006/jmla.1993.1038

Carreiras, M., Ferrand, L., Grainger, J., & Perea, M. (2005). Sequential effects of phonological priming in visual word recognition. *Psychological Science*, *16*, 585-589. doi:10.1111/j.1467-9280.2006.01821.x

Center for Speech Technology Research, University of Edinburgh (n.d.), *Unisyn Lexicon* [Data file]. Retrieved July 19, 2004, from www.cstr.ed.ac.uk/projects/unisyn

Chateau, D., & Jared, D. (2003). Spelling–sound consistency effects in disyllabic word naming. *Journal of Memory and Language*, *48*, 255-280. doi:10.1016/S0749-596X(02)00521-1

Chetail, F., & Content, A. (2012). The internal structure of chaos: Letter category determines visual word perceptual units. *Journal of Memory and Language, 68*, 371-388. doi:dx.doi.org/10.1016/j.jml.2012.07.004

Chetail, F., & Content, A. (2013). Segmentation of written words in French. *Language and Speech, 56*, 125-144. doi: 10.1177/0023830912442919

Chetail, F., & Content, A. (2014). What is the difference between OASIS and OPERA? Roughly five pixels. Orthographic structure biases the perceived length of letter strings. *Psychological Science*,…

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

Chetail, F., & Mathey, S. (2009). Syllabic priming in lexical decision and naming tasks: The syllable congruency effect re-examined in French. *Canadian Journal of Experimental Psychology*, *63*, 40-48. doi:10.1037/a0012944

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: L. Erlbaum Associates.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). San Diego, CA: Academic Press.

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

Conrad, M., Grainger, J., & Jacobs, A. M. (2007). Phonology as the source of syllable frequency effects in visual word recognition: Evidence from French. *Memory & Cognition*, *35*, 974-983. doi:10.3758/BF03193470

Content, A., & Peereman, R. (1992). Single and multiple process models of print to speech conversion. In J. Alegria, D. Holender, J. Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 351-375). Amsterdam: North-Holland.

Cutler, A., Butterfield, S., & Williams, J. N. (1987). The perceptual integrity of syllabic onsets. *Journal of Memory and Language*, *26*, 406-418. doi:10.1016/0749-596X(87)90099-4

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

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. doi:10.1037/0096-1523.32.3.535

De Cara, B., & Goswami, U. (2002). Similarity relations among spoken words: The special status of rimes in English. *Behavior Research Methods, Instruments, & Computers*, *34*, 416-423. doi:10.3758/BF03195470

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, *125*, 777-799.

Ferrand, L., & New, B. (2003). Syllabic length effects in visual word recognition and naming. *Acta Psychologica, 113*, 167-183. doi:10.1016/S0001-6918(03)00031-3

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A. et al. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*, 488–496. doi:10.3758/BRM.42.2.488

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, *35*, 263-279. doi: 10.1017/S0140525X11001841

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 104-115.

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *The Quarterly Journal of Experimental Psychology*, *39*, 211-251. doi:10.1080/14640748708401785

Gibson, E. J. (1965). Learning to read. *Annals of Dyslexia*, *15*, 32–47.

Gómez, P., Ratcliff, R., & Perea, M. (2008). The Overlap Model: A model of letter position coding. *Psychological Review*, *115*, 577-600. doi:10.1037/a0012667

Grainger, J. & Van Heuven, W. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *The mental lexicon* (pp. 1-24). New York: Nova Science Publishers.

Kessler, B., Treiman, R., & Mullennix, J. (2008). Feedback-consistency effects in single-word reading. In E. L. Grigorenko & A. J. Naples (Eds.), *Single-word reading: Behavioral and biological perspectives* (pp. 159–174). New York: Erlbaum.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono-and disyllabic words and nonwords. *Frontiers in Psychology*, *1*, 1. doi: 10.3389/fpsyg.2010.00174

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*, 643-650. doi:10.3758/BRM.42.3.643

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, *44*, 287-304. doi:10.3758/s13428-011-0118-4

Lange, M. (2000). *Processus d'accès à la pronunciation dans le traitement des mots écrits* [*Processes of access to the pronunciation during written word processing*]. Université Libre de Bruxelles, Brussels, Belgium.

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. doi:10.1016/j.jml.2009.01.002

Lee, C. H., & Taft, M. (2011). Subsyllabic structure reflected in letter confusability effects in Korean word recognition. *Psychonomic Bulletin & Review*, *18*, 129–134. doi:10.3758/s13423-010-0028-y

Levelt, W. J. ., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, *50*, 239–269. doi:10.1016/0010-0277(94)90030-2

Longtin, C.-M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *18*, 313–334. doi:10.1080/01690960244000036

Lupker, S. J., Acha, J., Davis, C. J., & Perea, M. (2012). An investigation of the role of grapheme units in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 1491-1516. doi:10.1037/a0026886

Maionchi-Pino, N., Magnan, A., & Ecalle, J. (2010). Syllable frequency effects in visual word recognition: Developmental approach in French children. *Journal of Applied Developmental Psychology*, *31*(1), 70–82. doi: 10.1016/j.appdev.2009.08.003

Marinus, E., & De Jong, P. F. (2011). Dyslexic and typical-reading children use vowel digraphs as perceptual units in reading. *The Quarterly Journal of Experimental Psychology*, *64*, 504-516. doi:10.1080/17470218.2010.509803

Massaro, D. W., Venezky, R. L., & Taylor, G. A. (1979). Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General, 108*, 107-124. doi:10.1037/0096-3445.108.1.107

Merriam-Webster's collegial dictionary (2013). Retrieved June, 8, 2013, from http://www.merriam-webster.com/.

Muncer, S. J., & Knight, D. C. (2012). The bigram trough hypothesis and the syllable number effect in lexical decision. *Quarterly Journal of Experimental Psychology*, *65*, 2221–2230. doi:10.1080/17470218.2012.697176

Muncer, S. J., Knight, D., & Adams, J. W. (in press). Bigram frequency, number of syllables and morphemes and their effects on lexical decision and word naming. *Journal of Psycholinguistic Research*. doi:10.1007/s10936-013-9252-8

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13*, 45–52. doi:10.3758/BF03193811

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. doi:dx.doi.org/10.1016/j.jml.2009.11.002

Patterson, K. E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 335-359). London: Erlbaum.

Peereman, R., Brand, M., & Rey, A. (2006). Letter-by-letter processing in the phonological conversion of multiletter graphemes: Searching for sounds in printed pseudowords. *Psychonomic Bulletin & Review*, *13*, 38–44. doi:10.3758/BF03193810

Perea, M., abu Mallouh, R., & Carreiras, M. (2010). The search for an input-coding scheme: Transposed-letter priming in Arabic. *Psychonomic Bulletin & Review*, *17*, 375-380. doi:10.3758/BF03196438

Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception & Psychophysics, 61*, 691-706. doi:10.3758/BF03205539

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114*, 273□315. doi:10.1037/0033-295X.114.2.273

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

Prinzmetal, W., Treiman, R., & Rho, S. H. (1986). How to see a reading unit. *Journal of Memory and Language*, *25*, 461–475. doi:10.1016/0749-596X(86)90038-0

Rastle, K., & Coltheart, M. (1998). Whammies and double whammies: The effect of length on nonword reading. *Psychonomic Bulletin & Review, 5*, 277–282. doi:10.3758/BF03212951

Rey, A., Ziegler, J. C., & Jacobs, A. M. (2000). Graphemes are perceptual reading units. *Cognition, 75*, 1–12. doi:10.1016/S0010-0277(99)00078-5

Rouibah, A., & Taft, M. (2001). The role of syllabic structure in French word recognition. *Memory & Cognition, 29*, 373-381. doi: 10.3758/BF03194932

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention and Performance, XII: The Psychology of Reading* (pp. 245-263). Hillsdale: Lawrence Erlbaum Associates.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, *23*, 383-404. doi:10.1016/S0022-5371(84)90270-6

Seymour, P. H. K., Aro, M., Erskine, J. M., & Network, collaboration with C. A. A. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174. doi:10.1348/000712603321661859

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science*, *8*, 411-416. doi:10.2307/40063226

Spinelli, E., Kandel, S., Guerassimovitch, H., & Ferrand, L. (2012). Graphemic cohesion effect in reading and writing complex graphemes. *Language and Cognitive Processes*, *27*, 770-791. doi:10.1080/01690965.2011.586534

Spoehr, K. T., & Smith, E. E. (1973). The role of syllables in perceptual processing. *Cognitive Psychology*, *5*, 71–89. doi:10.1016/0010-0285(73)90026-1

Sze, W. P., Liow, S. J. R., & Yap, M. J. (2014). The Chinese Lexicon Project: A repository of lexical decision behavioral responses for 2,500 Chinese characters. *Behavior Research Methods*, *46*, 263–273. doi:10.3758/s13428-013-0355-9

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). Morphological processing: The BOSS re-ermerges. In M. Coltheart (Ed.), *Attention and Performance XII* (p. 265-279). Hillsdale, NJ: Erlbaum.

Taft, M. (1991). *Reading and the mental lexicon*. Exeter: Lawrence Erlbaum Associates.

Taft, M. (1992). The body of the BOSS - Sublexical units in the lexical processing of polysyllabic words. *Journal of Experimental Psychology: Leaning, Memory, and Cognition*, *18*, 1004-1014.

Taft, M., & Krebs-Lazendic, L. (2013). The role of orthographic syllable structure in assigning letters to their position in visual word recognition. *Journal of Memory and Language*, *68*, 85-97. doi:10.1016/j.jml.2012.10.004

Taft, Marcus, 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. doi:10.1075/ml.2.1.04taf

Treiman R., & Chafetz, J. (1987). Are there onset- and rime-like units in printed words? In M. Coltheart (Ed.), *Attention and Performance XII* (pp. 282-297). Hillslade: Lawrence Erlbaum Associates.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language*, *25*, 476-491. doi:10.1016/0749-596X(86)90039-2

Treiman, R., Kessler, B., Knewasser, S., Tincoff, R., & Bowman, M. (2000). English speakers' sensitivity to phonotactic patterns. In M. B. Broe & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the lexicon* (pp. 269-282). Cambridge, England: Cambridge University Press.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, *124*, 107-136.

Vallée, N., Rousset, I., Boë, L.-J. (2001). Des lexiques aux syllabes des langues du monde. *Linx, 45*, 37-50. doi : 10.4000/linx.721

Van den Bosch, A., Content, A., Daelemans, W., & De Gelder, B. (1994). Measuring the complexity of writing systems. *Journal of Quantitative Linguistics*, *1*(3), 178-188. doi:10.1080/09296179408590015

Velan, H., & Frost, R. (2009). Letter-transposition effects are not universal: The impact of transposing letters in Hebrew. *Journal of Memory and Language*, *61*, 285-302. doi:10.1016/j.jml.2009.05.003

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. doi:10.3758/BF03196158

Yap, M., J. (2007). *Visual word recognition: Explorations of megastudies, multisyllabic words, and individual differences*. Unpublished doctoral dissertation, Washington University in Saint Louis, Saint Louis, MO, USA.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*, 502-529. doi:10.1016/j.jml.2009.02.001

Yap, M. J., Liow, S. J., Jalil, S. B., & Faizal, S. S. (2010). The Malay Lexicon Project: A database of lexical statistics for 9,592 words. *Behavior Research Methods*, *42*, 992-1003. doi: 10.3758/BRM.42.4.992.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart’s N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*, 971-979. doi:10.3738/PBR.15.5.971

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, *131*, 3. doi:10.1037/0033-2909.131.1.3

Stone, G. O., & Van Orden, G. C. (1994). Building a resonance framework for word recognition using design and system principles. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 1248–1268.

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers*, *28*, 504-515. doi:10.3758/BF03200539

Table 1. Descriptive statistics for the continuous predictors and dependent variables in the regression analyses on the ELP

	Mean	Median	Standard deviation	Minimum	Maximum
LDT-RT	-0.06	-0.10	0.41	-0.92	2.28
LDT-Accuracy	0.79	0.88	0.23	0.03	1.00
Naming-RT	-0.06	-0.14	0.44	-0.94	2.68
Naming-Accuracy	0.91	0.96	0.13	0.08	1.00
Word frequency (log)	1.89	1.81	0.82	0.30	5.27
Number of letters	6.70	6.00	1.55	2.00	14.00
Bigram frequency	1847.16	1796.52	713.12	66.25	5121.25
OLD	2.61	2.50	0.80	1.10	7.05
Number of syllables	2.39	2.00	0.64	2.00	6.00
FFO consistency	0.84	0.90	0.16	0.02	1.00
FFR consistency	0.54	0.54	0.20	0.01	1.00
FBO consistency	0.74	0.77	0.18	0.01	1.00
FBR consistency	0.54	0.54	0.20	0.00	1.00

Table 2. Correlations between the continuous predictors and the dependent variables in the regression analyses on the ELP

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. LDT-RT	—	-.72***	.76***	-.59***	-.68***	.37***	-.04**	.46***	.34***	-.05***	-.05**	.02	-.05***
2. LDT-Accuracy		—	-.65***	.73***	.64***	-.08***	.09***	-.19***	-.09***	.01	.02	.02†	.04**
3. Naming-RT			—	-.67***	-.59***	.46***	-.00	.54***	.43***	-.16***	-.11***	-.01	-.13***
4. Naming-Accuracy				—	.51***	-.15***	.02	-.23***	-.18***	.07***	.11***	-.01	.10***
5. Word frequency					—	-.23***	.06***	-.27***	-.18***	.01	-.01	.02†	-.06***
6. Number of letters						—	.17***	.84***	.65***	-.13***	-.00	-.02	.01
7. Bigram frequency							—	-.10***	.09***	-.07***	-.06***	.16***	. 11***
8. OLD								—	.69***	-.11***	-.05***	.01	-.08***
9. Number of syllables									—	-.08***	-.27***	.08***	-.02
10. FFO consistency										—	-.02	.21***	.12***
11. FFR consistency											—	.03*	.22***
12. FBO consistency												—	-.05***
13. FBR consistency													—

Notes. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 3. Reaction time (RT) and accuracy coefficients in steps 1 to 3 of the regression analyses on the ELP for naming and lexical decision tasks

Predictors	Naming		LDT	
	RT	Accuracy	RT	Accuracy
Step 1				
R^2	**.577*****	**.298*****	**.557*****	**.435*****
Step 2				
Word type	.050**	-.012†	.000	-.012
R^2	**.577*****	**.298*****	**.557*****	**.435*****
Δ R^2	**.001****	**.000**†	**.000**	**.000**
Step 3				
Word frequency	-.255***	.076***	-.300***	.175***
Number of letters	-.019***	.013***	-.029***	.042***
Bigram frequency	.041***	-.011***	.027***	-.012**
OLD	.203***	-.033***	.180***	-.079***
Number of syllables	.064***	-.000	.037***	.014*
FFO consistency	-.221***	.037***	.014	-.004
FFR consistency	-.113***	.048***	-.009	.007
FBO consistency	.085***	-.014	.058*	.022
FBR consistency	-.236***	.060***	-.118***	.064***
Word type	.035*	-.005	-.011	-.004
Word type x Word frequency	-.104***	.053***	-.077***	.056***
R^2	**.580*****	**.304**	**.558*****	**.437**
Δ R^2	**.003*****	**.006*****	**.001*****	**.002*****

Notes. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$. The results for first phoneme predictors are not reported. When only the first phoneme variables were entered, they explained 0.6% and 5.1% of the variance in latencies in the lexical decision task and the naming task respectively (0.5% and 0.3% on accuracy respectively). Bigram frequency was divided by 1,000 so that the size of the scale was similar to that of the other predictors. For the sake of clarity, the details of the other predictors at step 1 and 2 are not included.

Table 4. Mean RT and accuracy coefficients, and variance explained across 4x100 regression analyses on the ELP (LDT and naming task)

Predictors	Naming		LDT	
	RT	Accuracy	RT	Accuracy
Step 1				
R^2	**.571**	**.318**	**.568**	**.454**
Step 2				
Word type	.121	-.014	.021	-.014
R^2	**.581**	**.320**	**.568**	**.455**
ΔR^2	**.009**	**.001**	**.001**	**.001**
Step 3				
Word type	.122	-.014	.021	-.015
Word type x Word Freq.	-.110	.055	-.079	.055
R^2	**.589**	**.340**	**.573**	**.464**
ΔR^2	**.008**	**.020**	**.005**	**.008**

Table 5. Reaction time (RT) and accuracy coefficients in steps 1 to 3 of the regression analyses for lexical decision performance in the BLP database

Predictors	RT	Accuracy
Step 1		
R^2	**.572*****	**.533*****
Step 2		
Word type	.013	-.021
R^2	**.572*****	**.533*****
Δ R^2	**.000**	**.0001**
Step 3		
Word frequency	-.427***	.193***
Number of letters	-.079***	.070***
MBF	.030***	-.009***
OLD	.273***	-.112***
Word type	-.015	-.002
Word type x Lexical Freq.	-.105*	.096***
R^2	**.572*****	**.534*****
Δ R^2	**.0004***	**.001*****

Notes. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. The results for first phoneme predictor are not reported.

Table 6. Mean RT coefficients, accuracy coefficients, and variance explained across 4x100 regression analyses in the lexical decision task (BLP)

Predictors	RT	Accuracy
Step 1		
R^2	**.627**	**.594**
Step 2		
Word type	.042	-.039
R^2	**.630**	**.599**
ΔR^2	**.003**	**.005**
Step 3		
Word type	.041	-.037
Word type x Word Freq.	-.103	.097
R^2	**.639**	**.624**
ΔR^2	**.009**	**.025**

Table 7. Bigrams coding for hiatus patterns and graphemic cohesion

Bigram	Frequency of mapping onto one phoneme	Example	Frequency of mapping onto two phonemes	Example	Graphemic cohesion
ai	7565.56	*afraid*	9.38	*Zaire*	1
ee	13791.76	*degree*	3.22	*preempt*	1
oo	12312.63	*bedroom*	39.33	*coordination*	1
oa	1207.48	*approach*	17.03	*koala*	0.99
eu	292.88	*Europe*	29.06	*museum*	0.91
ea	21768.03	*season*	2896.42	*create*	0.88
eo	1597.03	*people*	243.33	*video*	0.87
ie	3820.37	*belief*	630.64	*client*	0.86
oe	1237.71	*canoe*	207.57	*poet*	0.86
ei	1565.9	*neither*	634.41	*reimburse*	0.71
ui	493.33	*circuit*	219.67	*genuine*	0.69
ao	24.85	*pharaoh*	16.75	*chaos*	0.6
oi	1121.99	*avoid*	3206.18	*heroin*	0.26
uo	1.01	*fluorescence*	21.2	*duo*	0.05
ia	0	-	991.88	*diagram*	0
io	0	-	1007.57	*biography*	0
iu	0	-	70.84	*triumph*	0
ua	0	-	664.56	*January*	0
ue	0	-	33.27	*duet*	0
yi	0	-	839.99	*flying*	0
yo	0	-	17.22	*embryo*	0

Notes. Values correspond to token frequencies. Frequencies were computed on the 40,481 words in the ELP. The set therefore included monosyllabic and polymorphemic words.

Figure captions

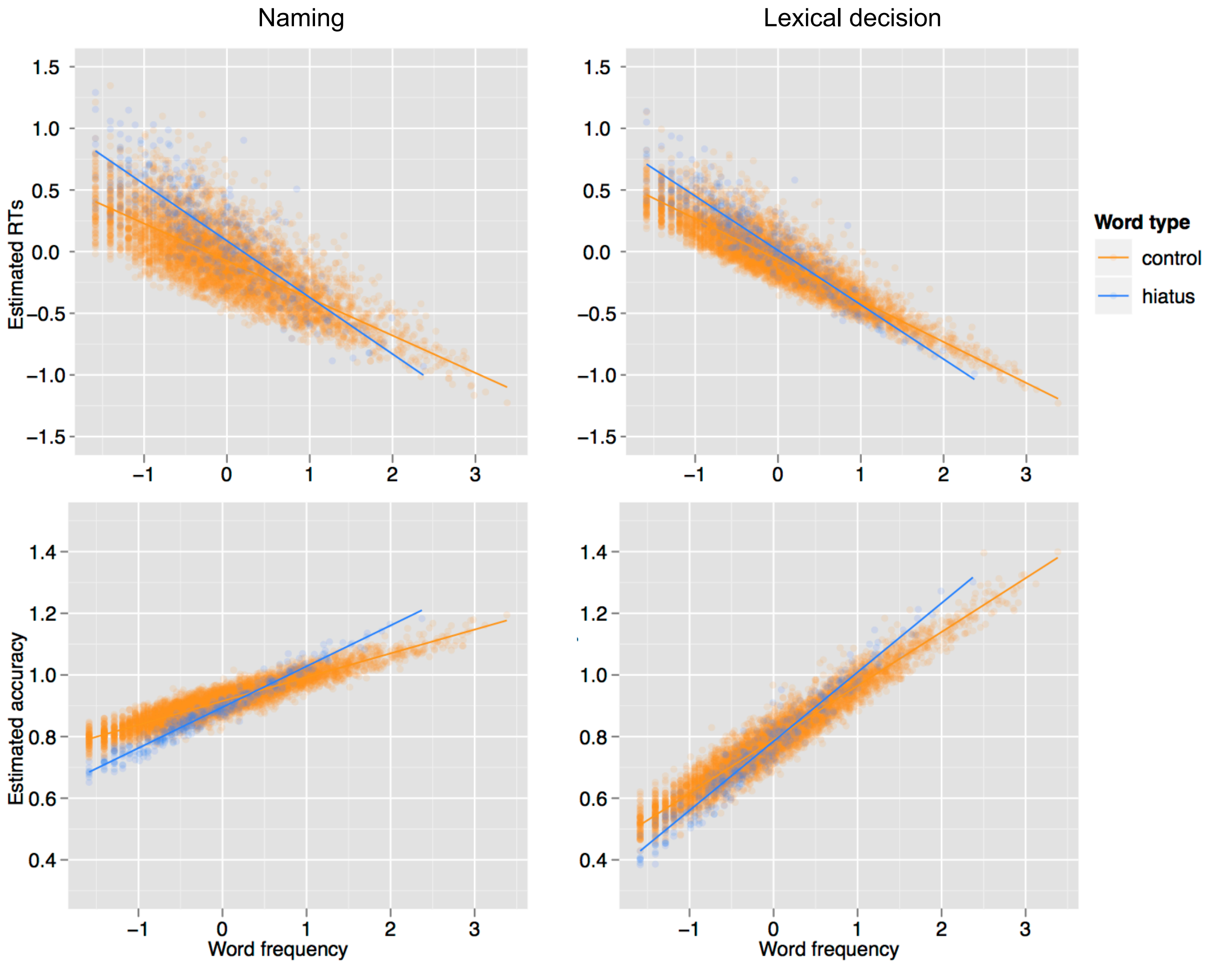
Figure 1. *Estimated RTs (upper panels) and accuracy (lower panels) at step 3 in the regression analysis as a function of word frequency in the naming task (left panels) and in the LDT (right panels) on the ELP*

Figure 2. *Percentage of regressions at Step 3 yielding a significant effect of word type (left panel) and of word type x word frequency interaction (right panel) for reaction times and accuracy in the lexical decision and naming tasks (ELP)*

Figure 3. *Percentage of regressions at Step 3 yielding a significant effect of word type (left panel) and of word type x word frequency interaction (right panel) on reaction times and accuracy in the LDT for the ELP and BLP common word subset)*

Figure 4. *Percentage of regressions leading to a significant word type effect in Step 3 in the LDT and naming tasks (ELP)*

Figure 1



Notes. Reaction times and accuracy are standardized, word frequency is log-transformed and centred.

Figure 2

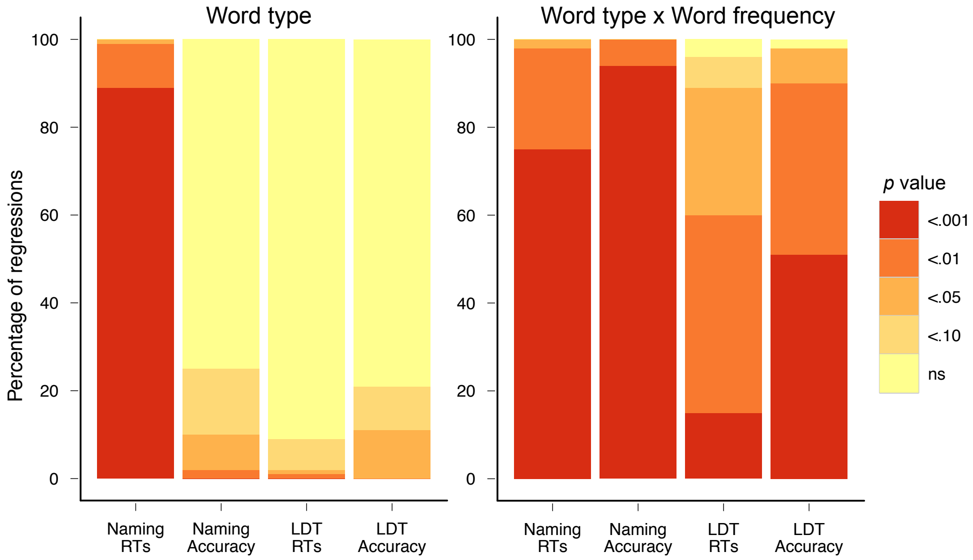

Word type
Word type x Word frequency
Percentage of regressions
100
80
60
40
20
0
Naming RTs
Naming Accuracy
LDT RTs
LDT Accuracy
p value
<.001
<.01
<.05
<.10
ns

Figure 3

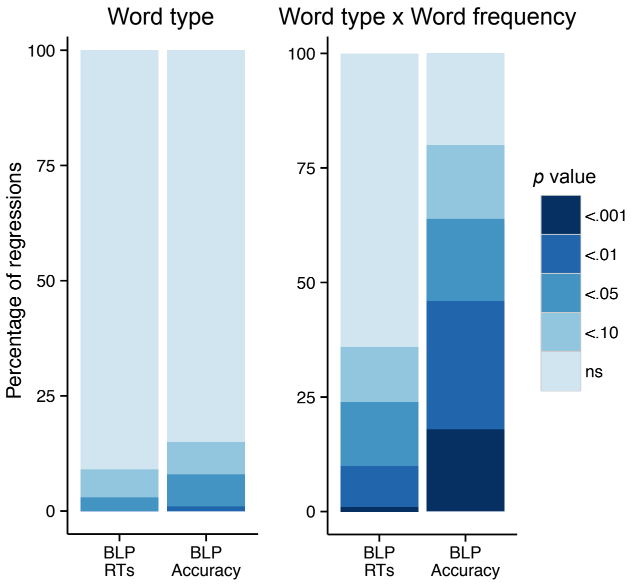
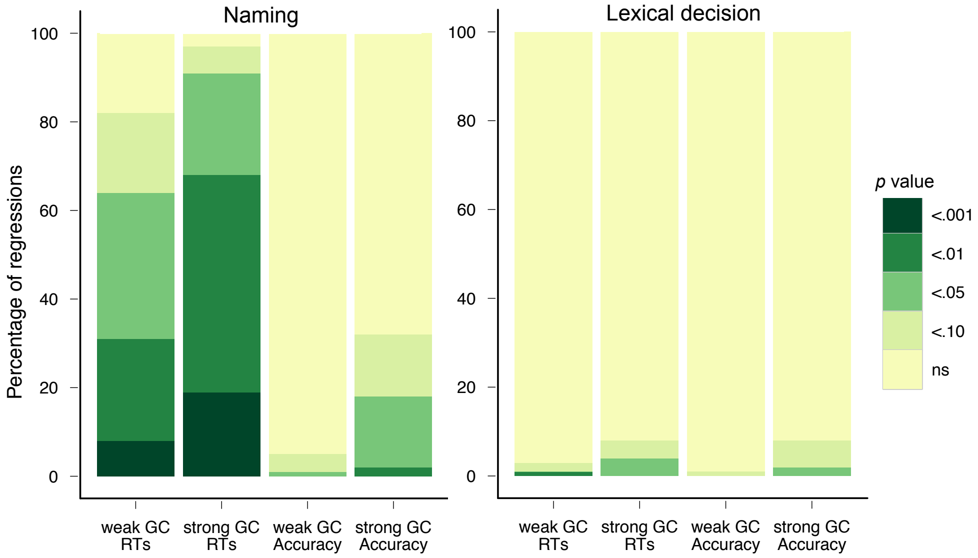

Word type
Word type x Word frequency
Percentage of regressions
100
75
50
25
0
BLP RTs
BLP Accuracy
p value
<.001
<.01
<.05
<.10
ns

Figure 4



Note. GC: graphemic cohesion.